

03806.0054-04.seq1st
SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: Blanc, Veronique
Blanche, Francis
Crouzet, Joel
Jacques, Nathalie
Lacroix, Patricia
Thibaut, Denis
Zagorec, Monique
Debussche, Laurent
De Crecy-Lagard, Valerie

(ii) TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
Coding For These Polypeptides And Their Use

(iii) NUMBER OF SEQUENCES: 43

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
(B) STREET: 1300 I Street, N.W., Suite 700
(C) CITY: Washington
(D) STATE: D.C.
(E) COUNTRY: USA
(F) ZIP: 20005-3315

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER: Unknown
(B) FILING DATE: 20-OCT-2003

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: US 09/635,359
(B) FILING DATE: 09-AUG-2000
(A) APPLICATION NUMBER: US 09/231,818
(B) FILING DATE: 15-JAN-1999
(A) APPLICATION NUMBER: US 08/403,852
(B) FILING DATE: 10-MAY-1995
(A) APPLICATION NUMBER: PCT/FR 93/00923
(B) FILING DATE: 25-SEP-1993

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: FR 92/11441
(B) FILING DATE: 25-SEP-1992

(viii) ATTORNEY/AGENT INFORMATION:

(A) NAME: Arrigo, Salvatore J.
(B) REGISTRATION NUMBER: 46,063
(C) REFERENCE/DOCKET NUMBER: 03806.0054-04000

(ix) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (202) 408-4000
(B) TELEFAX: (202) 408-4400

03806.0054-04.seq1st

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 5392 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S.pristinaespiralis*

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

GGATCCTGGC GTCCGCCGTC AAGAACTGAA CCGAGGAGAC ACCCACCATG ACCGCACCCC	60
GCCGGCGCAT CACCCTCGCC GGCATCATCG ACGGCCCCGG CGGCCATGTG GCCGCCTGGC	120
GCCACCCGGC GACCAAGGCG GACGCCCAGC TCGACTTCGA ATTCCACCGC GACAACGCCC	180
GCACCCTCGA ACGCGGCCTG TTCGACGCCG TGTTTCATCGC GGACATCGTC GCCGTGTGGG	240
GCACCCGCTT GGACTCCCTG TGCCGCACCT CGCGCACCGA GCACTTCGAA CCGCTCACCC	300
TGCTCGCCGC CTACGCCGCG GTCACCGAGC ACATCGGCCT GTGCGCCACC GCCACCACCA	360
CGTACAACGA ACCGGCGCAC ATCGCCGCCG GCTTCGCCTC CCTCGACCAC CTCAGCGGCG	420
GCCGGGCCGG CTGGAACGTC GTCACCTCCG CCGCACCGTG GGAGTCCGCC AACTTCGGCT	480
TCCCCGAGCA CCTGGAGCAC GGCAAACGCT ACGAGCGGGC CGAGGAGTTC ATCGACGTCG	540
TCAAAAAACT GTGGGACAGC GACGGCCGCC CCGTCGACCA CCGCGGCACC CACTTCGAGG	600
CCCCCGGCCC GCTCGGGATC GCCCGCCCCC CGCAGGGCCG CCCCCTCATC ATCCAGGCCG	660
GCTCCTCGCC GGTGGGACGC GAGTTCGCCG CCCGGCACGC CGAGGTCATC TTCACCCGGC	720
ACAACCGGCT CTCCGACGCC CAGGACTTCT ACGGCGACCT CAAGGCACGC GTCGCCCCGGC	780
ACGGCCGCGA CCCCAGAGAAG GTCCTCGTGT GGCCGACCCT CGCGCCGATC GTCGCCGCCA	840
CCGACACCGA GGCGAAGCAG CGCCTGCAGG AACTGCAGGA CCTCACCCAC GACCATGTCG	900
CCCTGCGCAC CCTTCAGGAC CACCTCGGCG ACGTCGACCT GAGCGCGTAC CCGATCGACG	960
GGCCCCGTCCC CGACATCCCG TACACCAACC AGTCCCAGTC GACGACCGAG CGGCTGATCG	1020
GCCTGGCCAG GCGCGAGAAC CTCAGCATCC GCGAGCTGGC CCTGCGGCTG ATGGGCGACA	1080
TCGTCTGTCG CACACCGGAG CAGCTCGCCG ACCACATGGA GAGCTGGTTC ACCGGCCGCG	1140
GCGCCGACGG CTTCAACATC GACTTCCCGT ACCTGCCGGG CTCCGCCGAC GACTTCGTCG	1200
ACCACGTGGT GCCCCAACTG CAGCGCCGCG GCCTGTACCG CTCGGGCTAC GAGGGCACCA	1260
CCCTGCGGGC CAACCTCGGC ATCGACGCCC CCCGGAAGGC AGGTGCAGCG GCTTGACTTC	1320

03806.0054-04.seq1st

CGTCCTAAAG	GCGGGGGATT	CCAGCGGTCG	CCCGCTGGGG	TTCCTGCTTC	ACCGACGACC	1380	
GCCCCGTCCG	GGAGGACTCC	CGTTGAGGTC	TTATACCGTC	TCCACAGGCC	GACGCCGCCA	1440	
GCCCCGGCGC	CAGGATGTTG	CGTGCCGCAT	TCACGTCGCG	GTCATGCACA	GCGCCGCAGT	1500	
CGCACGTCCA	CTCCCGGACG	TTCAGCGGCA	GCTTCCCGCG	GACCGTGCCG	CAGGTTCCGC	1560	
ACAGCTTGGA	GCTGGGGAAC	CAGCGGTCGA	TCACGACGAG	TTCGCGCCCA	TACCAGGCGC	1620	
ACTTGTA	CTC	CAGCATGGAG	CGCAGTTCCG	TCCAGGCCGC	GTCGGAGATG	GCGCGCGCGA	1680
GCTTGCCGTT	CTTCAGCAGG	TTGCGGACGG	TGAGGTCCTC	GATCACGACC	GTTTG	GTTTCT	1740
CACGGACGAG	TCGAGTCGAC	AGCTTG	TGGA	GGAAGTCGCA	GCGCCGGTCG	GTGATCCGGG	1800
CGTGGACGCG	GGCGACCTTG	CGGCGGGCTT	TCTTCCGGTT	CGCCGACCCC	TTCGCCTTGC		1860
GCGACACGTC	CCGCTGAGCC	TTGCGGAGGC	GGGCGCGGTC	ACGGCGCTCG	TGCTTGGGGT		1920
TGGTGATCTT	CTCCCCGGTG	GACAGGGTCA	CCAGGGAGGT	GATCCCGGCG	TCGATGCCGA		1980
CGGCCGCCGT	GGTGGCGGGC	GCGGGGGTGA	TGGTGTCCTC	GCACAGCAGG	GACACGAACC		2040
AGCGGCCCCG	ACGGTCGCGG	GACACGGTCA	CCGTCGTCGG	CTCCGCCCCT	TCGGGAAGGG		2100
GACGGGACCA	GCGGATGTCC	AGGGGCTCCG	CGGTCTTCGC	CAGCGTGAGC	TGTCCGTTAC		2160
GCCACGTGAA	GGCGCTGCGG	GTGTA	CTCGG	CCGACGCCCT	GGACTTTTTT	CGCGACTTGT	2220
ACCGCGGGTA	CTTCGACCGC	TTGGCGAAGA	AGTTGGCGAA	CGCCGTCTGC	AAGTGCCGCA		2280
GCGCCTGCTG	GAGCGGGACG	GAGGACACCT	CCGAGAGGAA	GGCGAGTTCT	TCGGTCTTCT		2340
TCCACTCCGT	CAGCGCGGCG	GACGACTGCA	CGTAGGAGAC	CCGGCGCTGC	TCGCCGTACC		2400
AGGCTCGCGT	GCGCCCCTCA	AGCGCCTTGT	TGTACACGAG	GCGGACACAG	CCGAACGTGC		2460
GGGACAGCTC	AGCCGCCTGC	TCGTCCGTGG	GATAAAAGCG	GTACTTGAAA	GCCCCGTTGA		2520
CCTGCTGCAT	CACGCCTCAC	ACGCTATCAG	TTCCCGTGTG	AGCGGCGGGT	GTCTGCCGGT		2580
GGTTGCAGAC	GCCGAACCGC	CCTGGCGGCG	ATTCGCCCAT	CCCTGCCCTG	CTCCGCAAGA		2640
GCTTCGTCTC	CTCCCCGGTC	TGAAGGCCGG	GGTATCCACG	AAGGAATTCT	GATGACCGCG		2700
CCCATCCTCG	TCGCCACCCT	CGACACCCGC	GGCCCCGCGG	CCACCCTCGG	CACGATCACC		2760
CGCGCCGTGC	GGGCCGCGGA	GGCCGCGGGA	TTGACGCGCG	TCCTGATCGA	CGACCGGGCC		2820
GCCGCCGGCG	TCCAGGGCCG	GTTGAGACG	ACGACGCTGA	CCGCCGCGCT	GGCCGCCGTC		2880
ACCGAGCACA	TCGGCCTGAT	CACCGCCCCG	CTCCCGGCGG	ACCAGGCCCC	CTACCACGTG		2940
TCCCGGATCA	CCGCCTCGCT	CGACCACCTC	GCCCACGGCC	GCACCGGCTG	GCTCGCGAGC		3000
ACGGACACCA	CCGACCCCGA	GGGCCGCACC	GGCGAACTCA	TCGACGTCGT	CCGCGGCCTG		3060
TGGGACAGCT	TCGACGACGA	CGCCTTCGTC	CACGACCGCG	CCGACGGCCT	GTACTGGCGG		3120
CTGCCCCGCG	TCCACCAACT	CGACCACCAG	GGCAGGCACT	TCGACGTGGC	CGGCCCCCTC		3180

03806.0054-04.seq1st

AACGTCGCCC	GCCCCGCGCA	GGGCCACCCC	GTCGTCGCCG	TCACCGGCCC	CGCCCTCGCC	3240
GCGGCCGCCG	ACCTCGTCCT	GCTCGACGAG	GCGGCCGACG	CCGCCTCGGT	GAAGCAGCAG	3300
GCACCGCACG	CCAAGATCCT	CCTGCCGCTG	CCCGGCCCGG	CCGCCGAAC	GCCCCGCGAC	3360
AGCCCCGCGG	ACGGCTTCAC	GGTGGCGCTC	ACCGGCTCCG	ACGACCCGGT	CCTGGCCGCG	3420
CTCGCCGCCC	GGCCCCGGCCG	CCCGGACCGC	ACCGCGGCCA	CCACCCTGCG	CGAACGCCTG	3480
GGCCTGGCCC	GCCCCGAGAG	CCGCCACGCC	CTCACCACCG	CCTGACGACC	CGTCCGCCCCG	3540
CTGCTTCCTG	GAGAGTCATG	TCCCGTCGCC	TGTTCACTC	GGAGTCCGTG	ACCGAGGGCC	3600
ACCCCGACAA	GATCGCCGAC	CAGATCAGTG	ACACCGTCCT	CGACGCCCTG	CTGCGCGAGG	3660
ACCCCGCCTC	ACGCGTCGCG	GTCGAGACCC	TGATCACCAC	CGGCCAGGTC	CACATCGCCG	3720
GCGAGGTCAC	CACCAAGGCG	TACGCGCCCA	TCGCCCAACT	GGTCCGCGAC	ACGATCCTGG	3780
CCATCGGCTA	CGACTCGTCC	GCCAAGGGCT	TCGACGGCGC	CTCCTGCGGC	GTCTCCGTCT	3840
CCATCGGCGC	GCAGTCCCCG	GACATCGCCC	AGGGCGTCGA	CAGCGCCTAC	GAGACCCGCG	3900
TCGAGGGCGA	GGACGACGAG	CTCGACCAGC	AGGGCGCCGG	CGACCAGGGC	CTGATGTTCTG	3960
GCTACGCCAC	CGACGAGACC	CCCTCGCTGA	TGCCGCTGCC	CATCGAGCTC	GCCCACCGCC	4020
TCTCGCGCCG	GCTCACCAG	GTCCGCAAGG	ACGGCACCGT	CCCCTACCTG	CGCCCCGACG	4080
GCAAGACCCA	GGTCACCATC	GAGTACCAGG	GCAGCCGCCC	GGTGCGCCTG	GACACCGTCG	4140
TCGTCTCCTC	CCAGCACGCC	GCCGACATCG	ACCTCGGCTC	CCTGCTCACC	CCCGACATCC	4200
GCGAGCACGT	CGTCGAGCAC	GTCTCGCCG	CACTCGCCGA	GGACGGCATC	AAGCTCGAGA	4260
CGGACAACTA	CCGCCTGCTG	GTCAACCCGA	CCGGCCGTTT	CGAGATCGGC	GGCCCGATGG	4320
GCGACGCCGG	CCTGACCGGC	CGCAAGATCA	TCATCGACAC	GTACGGCGGC	ATGGCCCGCC	4380
ACGGCGGTGG	CGCGTTCTCC	GGCAAGGACC	CGTCCAAGGT	CGACCGTTCC	GCCGCGTACG	4440
CGATGCGCTG	GGTCGCCAAG	AACGTCGTCG	CCGCGGGCCT	CGCCTCCCGC	TGCGAGGTCC	4500
AGGTCGCCTA	CGCCATCGGC	AAGGCCGAGC	CGGTCGGCCT	GTTCTGTCGAG	ACGTTCTGGCA	4560
CCGGCACCGT	CGCCAGGAG	CGCATCGAGA	AGGCCATCAC	CGAGGTCTTC	GACCTGCGCC	4620
CCGCGGCCAT	CATCCGCGAC	CTCGACCTGC	TGCGGCCCAT	CTACGCCGCC	ACCGCCGCCT	4680
ACGGCCACTT	CGGCCGCGAA	CTGCCCGACT	TCACCTGGGA	GCGGACCGAC	CGCGCCACC	4740
GGCTCAAGGC	CGCGGCCGGT	CTCTGAGCCG	GCCGGACCTG	TGAGGAGACC	TGACGTGCGC	4800
ATCGCTGTCA	CCGGTTCCAT	CGCCACCGAC	CATCTGATGG	TCTTCCCCGG	CCGGTTCGCG	4860
GATCAGCTGA	TCCCCGACCA	GCTCGCTCAT	GTCTCGCTCT	CCTTCCTGGT	CGACGCACTC	4920
GAGGTGCGCC	GGGGCGGAGT	GGCGGACAAC	GTCGCCTTCG	GCCTCGGCGG	CCTCGGCCTC	4980
ACCCCCCAGC	TGGTCGGCGC	CGTGGGCAGC	GACTTCGCCG	AGTACGAGGT	CTGGCTCAAG	5040
GAACACGGCG	TCGACACCGG	CCCCGTCCTG	GTCTCCACCG	AGCGGCAGAC	CGCCCGGTTC	5100

03806.0054-04.seq1st

```

ATGTGCATCA CCGACCAGGA CTCCAACCAG ATCGCCTCCT TCTACGCGGG CGCCATGCAA 5160
GAGGCCCGCG ACATCGACCT GTGGCACCTG ACCACCGGCA GCGTCCGCCC CGACCTCGTC 5220
CTGGTCTGCC CGAACGACCC GGC GGCGATG CTGCGCCACA CGGGGAGTGC CGCGAAACTG 5280
GGCCTGCCGT TCGCCGCCGA CCCCTCCCAG CAGCTCGCCC GCCTGGAGGG AGGGAGGTAC 5340
GCGAACTCGG TCGACGGGGC CCGTTGTTT TTCACCAACG AAGTACGAGG CC 5392

```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1268 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..1268

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```

ATG ACC GCA CCC CGC CGG CGC ATC ACC CTC GCC GGC ATC ATC GAC GGC 48
Met Thr Ala Pro Arg Arg Arg Ile Thr Leu Ala Gly Ile Ile Asp Gly
 1          5          10          15

CCC GGC GGC CAT GTG GCC GCC TGG CGC CAC CCG GCG ACC AAG GCG GAC 96
Pro Gly Gly His Val Ala Ala Trp Arg His Pro Ala Thr Lys Ala Asp
          20          25          30

GCC CAG CTC GAC TTC GAA TTC CAC CGC GAC AAC GCC CGC ACC CTC GAA 144
Ala Gln Leu Asp Phe Glu Phe His Arg Asp Asn Ala Arg Thr Leu Glu
          35          40          45

CGC GGC CTG TTC GAC GCC GTG TTC ATC GCG GAC ATC GTC GCC GTG TGG 192
Arg Gly Leu Phe Asp Ala Val Phe Ile Ala Asp Ile Val Ala Val Trp
          50          55          60

GGC ACC CGC CTG GAC TCC CTG TGC CGC ACC TCG CGC ACC GAG CAC TTC 240
Gly Thr Arg Leu Asp Ser Leu Cys Arg Thr Ser Arg Thr Glu His Phe
          65          70          75          80

GAA CCG CTC ACC CTG CTC GCC GCC TAC GCC GCG GTC ACC GAG CAC ATC 288
Glu Pro Leu Thr Leu Leu Ala Ala Tyr Ala Ala Val Thr Glu His Ile
          85          90          95

GGC CTG TGC GCC ACC GCC ACC ACC ACG TAC AAC GAA CCG GCG CAC ATC 336
Gly Leu Cys Ala Thr Ala Thr Thr Thr Tyr Asn Glu Pro Ala His Ile
          100          105          110

```

03806.0054-04.seq1st

GCC Ala	GCC Ala	CGC Arg 115	TTC Phe	GCC Ala	TCC Ser	CTC Leu	GAC Asp 120	CAC His	CTC Leu	AGC Ser	GGC Gly 125	GGC Gly 125	CGG Arg	GCC Ala	GGC Gly	384
TGG Trp 130	AAC Asn	GTC Val	GTC Val	ACC Thr	TCC Ser	GCC Ala 135	GCA Ala	CCG Pro	TGG Trp	GAG Glu 140	TCC Ser 140	GCC Ala	AAC Asn	TTC Phe	GGC Gly	432
TTC Phe 145	CCC Pro	GAG Glu	CAC His	CTG Leu	GAG Glu 150	CAC His	GGC Gly	AAA Lys	CGC Arg	TAC Tyr 155	GAG Glu	CGG Arg	GCC Ala	GAG Glu	GAG Glu 160	480
TTC Phe	ATC Ile	GAC Asp	GTC Val	GTC Val 165	AAA Lys	AAA Lys	CTG Leu	TGG Trp	GAC Asp 170	AGC Ser	GAC Asp	GGC Gly	CGC Arg	CCC Pro 175	GTC Val	528
GAC Asp	CAC His	CGC Arg	GGC Gly 180	ACC Thr	CAC His	TTC Phe	GAG Glu 185	GCC Ala 185	CCC Pro	GGC Gly	CCG Pro	CTC Leu	GGG Gly 190	ATC Ile	GCC Ala	576
CGC Arg	CCC Pro	CCG Pro 195	CAG Gln	GGC Gly	CGC Arg	CCC Pro	GTC Val 200	ATC Ile	ATC Ile	CAG Gln	GCC Ala 205	GGC Gly 205	TCC Ser	TCG Ser	CCG Pro	624
GTG Val 210	GGA Gly	CGC Arg	GAG Glu	TTC Phe	GCC Ala	GCC Ala 215	CGG Arg	CAC His	GCC Ala	GAG Glu 220	GTC Val 220	ATC Ile	TTC Phe	ACC Thr	CGG Arg	672
CAC His 225	AAC Asn	CGG Arg	CTC Leu	TCC Ser	GAC Asp 230	GCC Ala	CAG Gln	GAC Asp	TTC Phe	TAC Tyr 235	GGC Gly	GAC Asp	CTC Leu	AAG Lys	GCA Ala 240	720
CGC Arg	GTC Val	GCC Ala	CGG Arg	CAC His 245	GGC Gly	CGC Arg	GAC Asp	CCC Pro	GAG Glu 250	AAG Lys	GTC Val	CTC Leu	GTG Val	TGG Trp 255	CCG Pro	768
ACC Thr	CTC Leu	GCG Ala	CCG Pro 260	ATC Ile	GTC Val	GCC Ala	GCC Ala	ACC Thr 265	GAC Asp	ACC Thr	GAG Glu	GCG Ala	AAG Lys 270	CAG Gln	CGC Arg	816
CTG Leu	CAG Gln 275	GAA Glu	CTG Leu	CAG Gln	GAC Asp	CTC Leu	ACC Thr 280	CAC His	GAC Asp	CAT His	GTC Val	GCC Ala 285	CTG Leu	CGC Arg	ACC Thr	864
CTT Leu 290	CAG Gln	GAC Asp	CAC His	CTC Leu	GGC Gly	GAC Asp 295	GTC Val	GAC Asp	CTG Leu	AGC Ser	GCG Ala 300	TAC Tyr	CCG Pro	ATC Ile	GAC Asp	912
GGG Gly 305	CCC Pro	GTC Val	CCC Pro	GAC Asp	ATC Ile 310	CCG Pro	TAC Tyr	ACC Thr	AAC Asn	CAG Gln 315	TCC Ser	CAG Gln	TCG Ser	ACG Thr	ACC Thr 320	960
GAG Glu	CGG Arg	CTG Leu	ATC Ile	GGC Gly 325	CTG Leu	GCC Ala	AGG Arg	CGC Arg	GAG Glu 330	AAC Asn	CTC Leu	AGC Ser	ATC Ile	CGC Arg 335	GAG Glu	1008
CTG Leu	GCC Ala	CTG Leu	CGG Arg 340	CTG Leu	ATG Met	GGC Gly	GAC Asp	ATC Ile 345	GTC Val	GTC Val	GGC Gly	ACA Thr	CCG Pro 350	GAG Glu	CAG Gln	1056
CTC Leu	GCC Ala	GAC Asp 355	CAC His	ATG Met	GAG Glu	AGC Ser	TGG Trp 360	TTC Phe	ACC Thr	GGC Gly	CGC Arg	GGC Gly 365	GCC Ala	GAC Asp	GGC Gly	1104

03806.0054-04.seq1st

TTC AAC ATC GAC TTC CCG TAC CTG CCG GGC TCC GCC GAC GAC TTC GTC	1152
Phe Asn Ile Asp Phe Pro Tyr Leu Pro Gly Ser Ala Asp Asp Phe Val	
370 375 380	
GAC CAC GTG GTG CCC GAA CTG CAG CGC CGC GGC CTG TAC CGC TCG GGC	1200
Asp His Val Val Pro Glu Leu Gln Arg Arg Gly Leu Tyr Arg Ser Gly	
385 390 395 400	
TAC GAG GGC ACC ACC CTG CGG GCC AAC CTC GGC ATC GAC GCC CCC CGG	1248
Tyr Glu Gly Thr Thr Leu Arg Ala Asn Leu Gly Ile Asp Ala Pro Arg	
405 410 415	
AAG GCA GGT GCA GCG GCT TG	1268
Lys Ala Gly Ala Ala Ala	
420	

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 833 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..833

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

ATG ACC GCG CCC ATC CTC GTC GCC ACC CTC GAC ACC CGC GGC CCC GCC	48
Met Thr Ala Pro Ile Leu Val Ala Thr Leu Asp Thr Arg Gly Pro Ala	
1 5 10 15	
GCC ACC CTC GGC ACG ATC ACC CGC GCC GTG CGG GCC GCG GAG GCC GCC	96
Ala Thr Leu Gly Thr Ile Thr Arg Ala Val Arg Ala Ala Glu Ala Ala	
20 25 30	
GGA TTC GAC GCC GTC CTG ATC GAC GAC CGG GCC GCC GCC GGC GTC CAG	144
Gly Phe Asp Ala Val Leu Ile Asp Asp Arg Ala Ala Ala Gly Val Gln	
35 40 45	
GGC CGG TTC GAG ACG ACG ACG CTG ACC GCC GCG CTG GCC GCC GTC ACC	192
Gly Arg Phe Glu Thr Thr Thr Leu Thr Ala Ala Leu Ala Ala Val Thr	
50 55 60	
GAG CAC ATC GGC CTG ATC ACC GCC CCG CTC CCG GCC GAC CAG GCC CCC	240
Glu His Ile Gly Leu Ile Thr Ala Pro Leu Pro Ala Asp Gln Ala Pro	
65 70 75 80	
TAC CAC GTG TCC CGG ATC ACC GCC TCG CTC GAC CAC CTC GCC CAC GGC	288
Tyr His Val Ser Arg Ile Thr Ala Ser Leu Asp His Leu Ala His Gly	
Page 7	

03806.0054-04.seq1st

85	90	95	
CGC ACC GGC TGG CTC GCG AGC ACG GAC ACC ACC GAC CCC GAG GGC CGC Arg Thr Gly Trp Leu Ala Ser Thr Asp Thr Thr Asp Pro Glu Gly Arg	100 105	110	336
ACC GGC GAA CTC ATC GAC GTC GTC CGC GGC CTG TGG GAC AGC TTC GAC Thr Gly Glu Leu Ile Asp Val Val Arg Gly Leu Trp Asp Ser Phe Asp	115 120	125	384
GAC GAC GCC TTC GTC CAC GAC CGC GCC GAC GGC CTG TAC TGG CGG CTG Asp Asp Ala Phe Val His Asp Arg Ala Asp Gly Leu Tyr Trp Arg Leu	130 135	140	432
CCC GCC GTC CAC CAA CTC GAC CAC CAG GGC AGG CAC TTC GAC GTG GCC Pro Ala Val His Gln Leu Asp His Gln Gly Arg His Phe Asp Val Ala	145 150	155 160	480
GGC CCC CTC AAC GTC GCC CGC CCG CCG CAG GGC CAC CCC GTC GTC GCC Gly Pro Leu Asn Val Ala Arg Pro Pro Gln Gly His Pro Val Val Ala	165 170	175	528
GTC ACC GGC CCC GCC CTC GCC GCG GCC GCC GAC CTC GTC CTG CTC GAC Val Thr Gly Pro Ala Leu Ala Ala Ala Asp Leu Val Leu Leu Asp	180 185	190	576
GAG GCG GCC GAC GCC GCC TCG GTG AAG CAG CAG GCA CCG CAC GCC AAG Glu Ala Ala Asp Ala Ala Ser Val Lys Gln Gln Ala Pro His Ala Lys	195 200	205	624
ATC CTC CTG CCG CTG CCC GGC CCG GCC GCC GAA CTG CCC GCC GAC AGC Ile Leu Leu Pro Leu Pro Gly Pro Ala Ala Glu Leu Pro Ala Asp Ser	210 215	220	672
CCC GCG GAC GGC TTC ACG GTG GCG CTC ACC GGC TCC GAC GAC CCG GTC Pro Ala Asp Gly Phe Thr Val Ala Leu Thr Gly Ser Asp Asp Pro Val	225 230	235 240	720
CTG GCC GCG CTC GCC GCC CGG CCC GGC CGC CCG GAC CGC ACC GCG GCC Leu Ala Ala Leu Ala Ala Arg Pro Gly Arg Pro Asp Arg Thr Ala Ala	245 250	255	768
ACC ACC CTG CGC GAA CGC CTG GGC CTG GCC CGC CCC GAG AGC CGC CAC Thr Thr Leu Arg Glu Arg Leu Gly Leu Ala Arg Pro Glu Ser Arg His	260 265	270	816
GCC CTC ACC ACC GCC TG Ala Leu Thr Thr Ala	275		833

(2) INFORMATION FOR SEQ ID NO: 4:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1208 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..1208

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATG TCC CGT CGC CTG TTC ACC TCG GAG TCC GTG ACC GAG GGC CAC CCC	48
Met Ser Arg Arg Leu Phe Thr Ser Glu Ser Val Thr Glu Gly His Pro	
1 5 10 15	
GAC AAG ATC GCC GAC CAG ATC AGT GAC ACC GTC CTC GAC GCC CTG CTG	96
Asp Lys Ile Ala Asp Gln Ile Ser Asp Thr Val Leu Asp Ala Leu Leu	
20 25 30	
CGC GAG GAC CCC GCC TCA CGC GTC GCG GTC GAG ACC CTG ATC ACC ACC	144
Arg Glu Asp Pro Ala Ser Arg Val Ala Val Glu Thr Leu Ile Thr Thr	
35 40 45	
GGC CAG GTC CAC ATC GCC GGC GAG GTC ACC ACC AAG GCG TAC GCG CCC	192
Gly Gln Val His Ile Ala Gly Glu Val Thr Thr Lys Ala Tyr Ala Pro	
50 55 60	
ATC GCC CAA CTG GTC CGC GAC ACG ATC CTG GCC ATC GGC TAC GAC TCG	240
Ile Ala Gln Leu Val Arg Asp Thr Ile Leu Ala Ile Gly Tyr Asp Ser	
65 70 75 80	
TCC GCC AAG GGC TTC GAC GGC GCC TCC TGC GGC GTC TCC GTC TCC ATC	288
Ser Ala Lys Gly Phe Asp Gly Ala Ser Cys Gly Val Ser Val Ser Ile	
85 90 95	
GGC GCG CAG TCC CCG GAC ATC GCC CAG GGC GTC GAC AGC GCC TAC GAG	336
Gly Ala Gln Ser Pro Asp Ile Ala Gln Gly Val Asp Ser Ala Tyr Glu	
100 105 110	
ACC CGC GTC GAG GGC GAG GAC GAC GAG CTC GAC CAG CAG GGC GCC GGC	384
Thr Arg Val Glu Gly Glu Asp Asp Glu Leu Asp Gln Gln Gly Ala Gly	
115 120 125	
GAC CAG GGC CTG ATG TTC GGC TAC GCC ACC GAC GAG ACC CCC TCG CTG	432
Asp Gln Gly Leu Met Phe Gly Tyr Ala Thr Asp Glu Thr Pro Ser Leu	
130 135 140	
ATG CCG CTG CCC ATC GAG CTC GCC CAC CGC CTC TCG CGC CGG CTC ACC	480
Met Pro Leu Pro Ile Glu Leu Ala His Arg Leu Ser Arg Arg Leu Thr	
145 150 155 160	
GAG GTC CGC AAG GAC GGC ACC GTC CCC TAC CTG CGC CCC GAC GGC AAG	528
Glu Val Arg Lys Asp Gly Thr Val Pro Tyr Leu Arg Pro Asp Gly Lys	
165 170 175	
ACC CAG GTC ACC ATC GAG TAC CAG GGC AGC CGC CCG GTG CGC CTG GAC	576
Thr Gln Val Thr Ile Glu Tyr Gln Gly Ser Arg Pro Val Arg Leu Asp	
180 185 190	
ACC GTC GTC GTC TCC TCC CAG CAC GCC GCC GAC ATC GAC CTC GGC TCC	624
Thr Val Val Val Ser Ser Gln His Ala Ala Asp Ile Asp Leu Gly Ser	
195 200 205	

03806.0054-04.seq1st

CTG	CTC	ACC	CCC	GAC	ATC	CGC	GAG	CAC	GTC	GTC	GAG	CAC	GTC	CTC	GCC	672
Leu	Leu	Thr	Pro	Asp	Ile	Arg	Glu	His	Val	Val	Glu	His	Val	Leu	Ala	
	210					215					220					
GCA	CTC	GCC	GAG	GAC	GGC	ATC	AAG	CTC	GAG	ACG	GAC	AAC	TAC	CGC	CTG	720
Ala	Leu	Ala	Glu	Asp	Gly	Ile	Lys	Leu	Glu	Thr	Asp	Asn	Tyr	Arg	Leu	
	225				230					235					240	
CTG	GTC	AAC	CCG	ACC	GGC	CGT	TTC	GAG	ATC	GGC	GGC	CCG	ATG	GGC	GAC	768
Leu	Val	Asn	Pro	Thr	Gly	Arg	Phe	Glu	Ile	Gly	Gly	Pro	Met	Gly	Asp	
				245					250					255		
GCC	GGC	CTG	ACC	GGC	CGC	AAG	ATC	ATC	ATC	GAC	ACG	TAC	GGC	GGC	ATG	816
Ala	Gly	Leu	Thr	Gly	Arg	Lys	Ile	Ile	Ile	Asp	Thr	Tyr	Gly	Gly	Met	
			260				265						270			
GCC	CGC	CAC	GGC	GGT	GGC	GCG	TTC	TCC	GGC	AAG	GAC	CCG	TCC	AAG	GTC	864
Ala	Arg	His	Gly	Gly	Gly	Ala	Phe	Ser	Gly	Lys	Asp	Pro	Ser	Lys	Val	
		275					280					285				
GAC	CGT	TCC	GCC	GCG	TAC	GCG	ATG	CGC	TGG	GTC	GCC	AAG	AAC	GTC	GTC	912
Asp	Arg	Ser	Ala	Ala	Tyr	Ala	Met	Arg	Trp	Val	Ala	Lys	Asn	Val	Val	
	290					295					300					
GCC	GCG	GGC	CTC	GCC	TCC	CGC	TGC	GAG	GTC	CAG	GTC	GCC	TAC	GCC	ATC	960
Ala	Ala	Gly	Leu	Ala	Ser	Arg	Cys	Glu	Val	Gln	Val	Ala	Tyr	Ala	Ile	
	305				310					315					320	
GGC	AAG	GCC	GAG	CCG	GTC	GGC	CTG	TTC	GTC	GAG	ACG	TTC	GGC	ACC	GGC	1008
Gly	Lys	Ala	Glu	Pro	Val	Gly	Leu	Phe	Val	Glu	Thr	Phe	Gly	Thr	Gly	
				325					330					335		
ACC	GTC	GCC	CAG	GAG	CGC	ATC	GAG	AAG	GCC	ATC	ACC	GAG	GTC	TTC	GAC	1056
Thr	Val	Ala	Gln	Glu	Arg	Ile	Glu	Lys	Ala	Ile	Thr	Glu	Val	Phe	Asp	
			340					345					350			
CTG	CGC	CCC	GCG	GCC	ATC	ATC	CGC	GAC	CTC	GAC	CTG	CTG	CGG	CCC	ATC	1104
Leu	Arg	Pro	Ala	Ala	Ile	Ile	Arg	Asp	Leu	Asp	Leu	Leu	Arg	Pro	Ile	
		355					360						365			
TAC	GCC	GCC	ACC	GCC	GCC	TAC	GGC	CAC	TTC	GGC	CGC	GAA	CTG	CCC	GAC	1152
Tyr	Ala	Ala	Thr	Ala	Ala	Tyr	Gly	His	Phe	Gly	Arg	Glu	Leu	Pro	Asp	
	370					375					380					
TTC	ACC	TGG	GAG	CGG	ACC	GAC	CGC	GCC	CAC	CGG	CTC	AAG	GCC	GCG	GCC	1200
Phe	Thr	Trp	Glu	Arg	Thr	Asp	Arg	Ala	His	Arg	Leu	Lys	Ala	Ala	Ala	
	385				390					395					400	
GGT	CTC	TG														1208
Gly	Leu															

(2) INFORMATION FOR SEQ ID NO: 5:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 1879 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 110..1858

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

GATCGGCTCC TGACGGAGCG GCGGCGCGCG GCGCGGGCGC ATCAGCGGCG TGTCAACGGC	60
GCTGCCGACA CTGGGCGCGA CGCGAGGACG AAGCCGGAAA GGACCAACG ATG CTG	115
Met Leu	
1	
GAC GGA TGC GTT CCC TGG CCC GAG GAT GTG GCC GCG AAG TAC CGG GCG	163
Asp Gly Cys Val Pro Trp Pro Glu Asp Val Ala Ala Lys Tyr Arg Ala	
5 10 15	
GCC GGC TAC TGG CGG GGC GAG CCG CTG GGC ATG CTG CTG GGC CGC TGG	211
Ala Gly Tyr Trp Arg Gly Glu Pro Leu Gly Met Leu Leu Gly Arg Trp	
20 25 30	
GCG GAG CAG TAC GGC GAG CGG GAG GCG CTG GTC GGC GCG GAC GGG TGC	259
Ala Glu Gln Tyr Gly Glu Arg Glu Ala Leu Val Gly Ala Asp Gly Cys	
35 40 45 50	
TCC CGT GTC ACC TAC CGT GCC CTG GAC CGC TGG TGC GAC CGG CTG GCG	307
Ser Arg Val Thr Tyr Arg Ala Leu Asp Arg Trp Cys Asp Arg Leu Ala	
55 60 65	
GCG GGG TTC GCG GCG CGC GGG ATC GGC GCC GGC GAG CGG GTG CTG GTG	355
Ala Gly Phe Ala Ala Arg Gly Ile Gly Ala Gly Glu Arg Val Leu Val	
70 75 80	
CAG CTG CCG AAC ACG CCC GAG TTC GTC GCG GTG TGC TTC GCG CTG TTC	403
Gln Leu Pro Asn Thr Pro Glu Phe Val Ala Val Cys Phe Ala Leu Phe	
85 90 95	
CGT CTG GGC GCG CTG CCG GTG TTC GCG CTG CCC GCG CAC CGT GCC GCC	451
Arg Leu Gly Ala Leu Pro Val Phe Ala Leu Pro Ala His Arg Ala Ala	
100 105 110	
GAG GTG GGG CAC CTG CTC GAG CTG TCC GGC GCC GTC GCC CAC ATC CTG	499
Glu Val Gly His Leu Leu Glu Leu Ser Gly Ala Val Ala His Ile Leu	
115 120 125 130	
CCG GGC ACC GGC ACC GGC TAC GAC CAT GTC GCG GCG GCC GTG GAG GCC	547
Pro Gly Thr Gly Thr Gly Tyr Asp His Val Ala Ala Ala Val Glu Ala	
135 140 145	
CGT GCC CGC CGT GCC CGC CCG GTG CAG GTG TTC GTG GCG GGC GAG GCG	595
Arg Ala Arg Arg Ala Arg Pro Val Gln Val Phe Val Ala Gly Glu Ala	
150 155 160	
CCC GCG GTG CTG CCC GAG GGG TTC ACC GCG CTG GCC GAC GTG GAC GGC	643
Pro Ala Val Leu Pro Glu Gly Phe Thr Ala Leu Ala Asp Val Asp Gly	
165 170 175	

03806.0054-04.seqlst

GAC Asp 180	CCG Pro 180	GTG Val 180	GCG Ala 180	CCG Pro 180	GCG Ala 185	GAC Asp 185	GTG Val 185	GAC Asp 185	GCC Ala 185	TTC Phe 190	CGA Arg 190	CGT Arg 190	GGC Gly 190	GTC Val 190	TTC Phe 190	691
CTG Leu 195	CTG Leu 195	TCC Ser 195	GGG Gly 200	GGG Gly 200	ACG Thr 200	ACC Thr 200	GCG Ala 200	CTG Leu 200	CCG Pro 205	AAG Lys 205	CTG Leu 205	ATC Ile 205	CCG Pro 205	CGC Arg 210	ACC Thr 210	739
CAC His 215	GAC Asp 215	GAC Asp 215	TAC Tyr 215	GCC Ala 215	TAC Tyr 215	CAG Gln 215	TGC Cys 220	CGG Arg 220	GTC Val 220	ACG Thr 220	GCC Ala 220	GGT Gly 225	ATC Ile 225	TGC Cys 225	GGC Gly 225	787
CTG Leu 230	GAC Asp 230	GCG Ala 230	GAC Asp 230	AGT Ser 230	GTC Val 230	TAT Tyr 230	CTG Leu 235	GCG Ala 235	GTG Val 235	CTG Leu 235	CCG Pro 240	GCC Ala 240	GAG Glu 240	TTC Phe 240	AAC Asn 240	835
TTC Phe 245	CCC Pro 245	TTC Phe 245	GGC Gly 250	TGC Cys 250	CCG Pro 250	GGC Gly 250	ATC Ile 250	CTG Leu 250	GGC Gly 250	ACC Thr 250	CTG Leu 255	CAC His 255	GCC Ala 255	GGC Gly 255	GGG Gly 255	883
CGG Arg 260	GTG Val 260	GTG Val 260	TTC Phe 260	GCG Ala 265	CTG Leu 265	TCA Ser 265	CCG Pro 265	CAG Gln 265	CCC Pro 270	GAG Glu 270	GAG Glu 270	TGC Cys 270	TTC Phe 270	GCG Ala 270	CTG Leu 270	931
ATC Ile 275	GAA Glu 275	CGC Arg 275	GAA Glu 275	CAC His 280	GTC Val 280	ACC Thr 280	TTC Phe 280	ACC Thr 280	TCC Ser 285	GTC Val 285	ATC Ile 285	CCC Pro 285	ACG Thr 285	ATC Ile 285	GTG Val 290	979
CAC His 295	CTG Leu 295	TGG Trp 295	CTG Leu 295	GCG Ala 295	GCC Ala 295	GCC Ala 295	GCA Ala 300	CAA Gln 300	GGC Gly 300	CAC His 300	GGC Gly 300	CGC Arg 300	GAC Asp 305	CTG Leu 305	GGC Gly 305	1027
AGC Ser 310	CTT Leu 310	CAG Gln 310	CTG Leu 310	CTG Leu 310	CAG Gln 310	GTC Val 310	GGC Gly 310	AGC Ser 315	GCC Ala 315	AAA Lys 315	CTC Leu 315	CAC His 315	GAG Glu 320	GAG Glu 320	CTC Leu 320	1075
GCC Ala 325	GCC Ala 325	CGG Arg 325	ATC Ile 325	GGC Gly 325	CCC Pro 325	GAA Glu 330	CTG Leu 330	GGG Gly 330	GTG Val 330	CGG Arg 330	CTG Leu 335	CAG Gln 335	CAG Gln 335	GTG Val 335	TTC Phe 335	1123
GGC Gly 340	ATG Met 340	GCC Ala 340	GAG Glu 340	GGA Gly 340	CTG Leu 345	CTG Leu 345	ACC Thr 345	TTC Phe 345	ACC Thr 345	CGC Arg 350	GAC Asp 350	GAC Asp 350	GAC Asp 350	CCG Pro 350	GCG Ala 350	1171
GAC Asp 355	GTG Val 355	GTG Val 355	CTG Leu 355	CGC Arg 360	ACC Thr 360	CAG Gln 360	GGC Gly 360	CGG Arg 365	CCG Pro 365	GTG Val 365	TCC Ser 365	GAG Glu 365	GCC Ala 365	GAC Asp 370	GAG Glu 370	1219
ATA Ile 375	CGC Arg 375	GTC Val 375	GCC Ala 375	GAC Asp 375	CCC Pro 375	GAC Asp 375	GGC Gly 375	CGG Arg 380	CCC Pro 380	GTG Val 380	CCC Pro 380	CGC Arg 380	GGT Gly 385	GAG Glu 385	ACC Thr 385	1267
GGT Gly 390	GAA Glu 390	CTG Leu 390	CTC Leu 390	ACC Thr 390	CGC Arg 390	GGC Gly 390	CCC Pro 395	TAC Tyr 395	ACG Thr 395	CTG Leu 395	CGC Arg 395	GGC Gly 400	TAC Tyr 400	TAC Tyr 400	CGG Arg 400	1315
GCC Ala 405	CCC Pro 405	GAG Glu 405	CAC His 405	AAC Asn 405	GCC Ala 410	CGC Arg 410	GCG Ala 410	TTC Phe 410	ACC Thr 410	GAG Glu 415	GAC Asp 415	GGC Gly 415	TTC Phe 415	TAC Tyr 415	CGC Arg 415	1363
AGC Ser 415	GGC Gly 415	GAT Asp 415	CTG Leu 415	GTG Val 415	CGG Arg 415	CTC Leu 415	ACC Thr 415	GCC Ala 415	GAC Asp 415	GGG Gly 415	CAG Gln 415	TTG Leu 415	GTG Val 415	GTG Val 415	GAG Glu 415	1411

03806.0054-04.seq1st

420	425	430	
GGC AGG ATC AAG GAC GTC GTC ATC CGC GGC GGC GAC AAG GTC TCC GCG Gly Arg Ile Lys Asp Val Val Ile Arg Gly Gly Asp Lys Val Ser Ala 435 440 445 450			1459
ACC GAG GTC GAG GGC CAC CTG GGC GCC CAC CCC GAC GTC CAG CAG GCC Thr Glu Val Glu Gly His Leu Gly Ala His Pro Asp Val Gln Gln Ala 455 460 465			1507
GCC GTC GTC GCC ATG CCC GAC CCG GTG TGG GGC GAG AAG GTC TGC GCC Ala Val Val Ala Met Pro Asp Pro Val Trp Gly Glu Lys Val Cys Ala 470 475 480			1555
TAC ATC GTG CCC GCA CCC GGC CGT CCC GCA CCG CCG ATG GCG GCG CTG Tyr Ile Val Pro Ala Pro Gly Arg Pro Ala Pro Pro Met Ala Ala Leu 485 490 495			1603
CGC CGG CTG CTG CGC GCG CGG GGA CTG GCC GAC TAC AAG CTT CCC GAC Arg Arg Leu Leu Arg Ala Arg Gly Leu Ala Asp Tyr Lys Leu Pro Asp 500 505 510			1651
CGG GTG GAG GTC GTC GAC GCG TTC CCG CTG ACC GGC CTC AAC AAG GTC Arg Val Glu Val Val Asp Ala Phe Pro Leu Thr Gly Leu Asn Lys Val 515 520 525 530			1699
GAC AAG AAG GCC CTG GCG GCC GAC ATC GCC GCC AAG ACC GCC CCC ACC Asp Lys Lys Ala Leu Ala Ala Asp Ile Ala Ala Lys Thr Ala Pro Thr 535 540 545			1747
CGC CCC ACC ACC GCC GGC CAC GGC CCG ACC ACG GAC GGC GAT ACG GCC Arg Pro Thr Thr Ala Gly His Gly Pro Thr Thr Asp Gly Asp Thr Ala 550 555 560			1795
GGT GGG GGT GGG TCC GCG GGC GGG GTG ACG GCC GCC GGT GGC GGG CGG Gly Gly Gly Gly Ser Ala Gly Gly Val Thr Ala Ala Gly Gly Gly Arg 565 570 575			1843
GAG GAG GCG GCG TGAGCGGGCC CGGGCCCGAG GGCG Glu Glu Ala Ala 580			1879

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1833 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 103..1689

03806.0054-04.seq1st

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

GGATCCCCTC	CCCCAGGGCC	CTGGCGGGCC	CGCCGGGCGG	TGGGGGAGGT	GCGGGGGCCG	60
CGGGCCCCGG	CACCGCACGA	ACAGAACAAC	CGCTCCGGGC	CC ATG CGG ACT TCA	114	
	Met	Arg	Thr	Ser		
	1					
CGG TCC CAC GAC CAG CGG GCC CCT ACC CCC TGG AGA CAT CCC TTG CAC	162					
Arg Ser His Asp Gln Arg Ala Pro Thr Pro Trp Arg His Pro Leu His						
5 10 15 20						
AGC ACC CGG CCC GCG CCC GCG GCC GAC CGT GAC CCC AGG CGC TGG GTC	210					
Ser Thr Arg Pro Ala Pro Ala Ala Asp Arg Asp Pro Arg Arg Trp Val						
25 30 35						
ATC CTC GGC GTG ATC TGC CTG GCC CAA CTC GTC GTC CTG CTC GAC AAC	258					
Ile Leu Gly Val Ile Cys Leu Ala Gln Leu Val Val Leu Leu Asp Asn						
40 45 50						
ACC GTC CTC AAC GTC GCC ATC CCG GTG CTC ACC ACC GAC CTG GGC GCC	306					
Thr Val Leu Asn Val Ala Ile Pro Val Leu Thr Thr Asp Leu Gly Ala						
55 60 65						
AGC ACC GCC GAC ATC CAG TGG ATG ATC AAC GCC TAC GCG CTC GTG CAG	354					
Ser Thr Ala Asp Ile Gln Trp Met Ile Asn Ala Tyr Ala Leu Val Gln						
70 75 80						
TCC GGG CTG CTG CTC ACC GCG GGC AGC CTC GCG GAC CGC TAC GGC CGC	402					
Ser Gly Leu Leu Leu Thr Ala Gly Ser Leu Ala Asp Arg Tyr Gly Arg						
85 90 95 100						
AAA CGG CTG CTG ATG CTC GGA CTG GTG CTC TTC GGC GCC GGG TCC GCC	450					
Lys Arg Leu Leu Met Leu Gly Leu Val Leu Phe Gly Ala Gly Ser Ala						
105 110 115						
TGG GCG GCC TTC GCC CAG GAC TCC GCC CAA CTC ATC GCC GCC CGG GCC	498					
Trp Ala Ala Phe Ala Gln Asp Ser Ala Gln Leu Ile Ala Ala Arg Ala						
120 125 130						
GGC ATG GGC GTG GGC GGG GCG CTG CTG GCG ACC ACC ACC CTC GCC GTC	546					
Gly Met Gly Val Gly Gly Ala Leu Leu Ala Thr Thr Thr Leu Ala Val						
135 140 145						
ATC ATG CAG GTC TTC GAC GAC GAC GAA CGC CCC CGG GCG ATC GGC CTG	594					
Ile Met Gln Val Phe Asp Asp Asp Glu Arg Pro Arg Ala Ile Gly Leu						
150 155 160						
TGG GGA GCG GCC AGC TCA CTG GGC TTC GCG GCC GGC CCG CTG CTC GGC	642					
Trp Gly Ala Ala Ser Ser Leu Gly Phe Ala Ala Gly Pro Leu Leu Gly						
165 170 175 180						
GGC GCC CTC CTC GAC CAC TTC TGG TGG GGC TCC ATC TTC CTG ATC AAC	690					
Gly Ala Leu Leu Asp His Phe Trp Trp Gly Ser Ile Phe Leu Ile Asn						
185 190 195						
CTG CCC GTC GCG CTG CTG GGC CTG CTG GCC GTC GCC CGC CTG GTG CCC	738					
Leu Pro Val Ala Leu Leu Gly Leu Leu Ala Val Ala Arg Leu Val Pro						
200 205 210						
GAG ACG AAG AAC CCC GAA GGC CGG CGC CCC GAC CTG CTC GGC GCC GTG	786					
Glu Thr Lys Asn Pro Glu Gly Arg Arg Pro Asp Leu Leu Gly Ala Val						

03806.0054-04.seq1st

215			220			225												
CTC	TCC	ACC	CTC	GGC	ATG	GTC	GGC	GTC	GTC	TAC	GCC	ATC	ATC	TCC	GGC			834
Leu	Ser	Thr	Leu	Gly	Met	Val	Gly	Val	Val	Tyr	Ala	Ile	Ile	Ser	Gly			
	230					235					240							
CCC	GAA	CAC	GGC	TGG	ACG	GCC	CCG	CAG	GTC	CTC	CTG	CCG	GCC	GCC	GTC			882
Pro	Glu	His	Gly	Trp	Thr	Ala	Pro	Gln	Val	Leu	Leu	Pro	Ala	Ala	Val			
245					250					255					260			
GCG	GCC	GCC	GCG	CTC	ACC	GCG	TTC	GTC	CGC	TGG	GAA	CTG	CAC	ACC	CCC			930
Ala	Ala	Ala	Ala	Leu	Thr	Ala	Phe	Val	Arg	Trp	Glu	Leu	His	Thr	Pro			
				265					270					275				
CAC	CCC	ATG	CTC	GAC	ATG	GGC	TTC	TTC	ACC	GAC	CGG	CGC	TTC	AAC	GGG			978
His	Pro	Met	Leu	Asp	Met	Gly	Phe	Phe	Thr	Asp	Arg	Arg	Phe	Asn	Gly			
			280					285					290					
CCG	TCG	CCG	GCG	GAG	TGC	TCG	TCG	TTC	GGC	ATG	GCC	GGC	TCG	CTC	TTC			1026
Pro	Ser	Pro	Ala	Glu	Cys	Ser	Ser	Phe	Gly	Met	Ala	Gly	Ser	Leu	Phe			
		295					300					305						
CTG	CTC	ACC	CAG	CAC	CTC	CAA	CTC	GTC	CTC	GGC	TAC	GAC	GCC	CTG	CAG			1074
Leu	Leu	Thr	Gln	His	Leu	Gln	Leu	Val	Leu	Gly	Tyr	Asp	Ala	Leu	Gln			
	310					315					320							
GCC	GGC	CTG	CGC	ACC	GCG	CCA	CTG	GCT	TTG	ACG	ATC	GTC	GCC	CTC	AAC			1122
Ala	Gly	Leu	Arg	Thr	Ala	Pro	Leu	Ala	Leu	Thr	Ile	Val	Ala	Leu	Asn			
325					330					335					340			
CTG	GCC	GGC	CTC	GGC	GCG	AAA	CTC	CTC	GCC	GCG	CTC	GGC	ACC	GCC	CGC			1170
Leu	Ala	Gly	Leu	Gly	Ala	Lys	Leu	Leu	Ala	Ala	Leu	Gly	Thr	Ala	Arg			
				345					350					355				
AGC	ATC	GCC	CTG	GGC	ATG	ACA	CTG	CTG	GCC	GCC	GGC	CTC	AGC	GCG	GTG			1218
Ser	Ile	Ala	Leu	Gly	Met	Thr	Leu	Leu	Ala	Ala	Gly	Leu	Ser	Ala	Val			
			360					365					370					
GCC	GTC	GGC	GGA	TCG	GGC	CCC	GAC	GCC	GGC	TAC	GGC	GGC	ATG	CTC	GCC			1266
Ala	Val	Gly	Gly	Ser	Gly	Pro	Asp	Ala	Gly	Tyr	Gly	Gly	Met	Leu	Ala			
		375					380					385						
GGC	CTG	CTC	CTC	ATG	GGC	GCG	GGC	ATC	GCA	CTG	GCC	ATG	CCC	GCC	ATG			1314
Gly	Leu	Leu	Leu	Met	Gly	Ala	Gly	Ile	Ala	Leu	Ala	Met	Pro	Ala	Met			
	390					395					400							
GCC	ACC	GCC	GTG	ATG	TCC	TCC	ATC	CCG	CCC	GCC	AAG	GCC	GGG	GCC	GGA			1362
Ala	Thr	Ala	Val	Met	Ser	Ser	Ile	Pro	Pro	Ala	Lys	Ala	Gly	Ala	Gly			
405					410					415					420			
GCG	GGC	GTG	CAG	GGC	ACC	CTG	ACC	GAG	TTC	GGC	GGC	GGA	CTG	GGA	GTG			1410
Ala	Gly	Val	Gln	Gly	Thr	Leu	Thr	Glu	Phe	Gly	Gly	Gly	Leu	Gly	Val			
				425					430				435					
GCG	ATC	CTC	GGC	GCC	GTC	CTC	GGC	TCC	CGC	TTC	GCC	TCC	CAA	CTG	CCC			1458
Ala	Ile	Leu	Gly	Ala	Val	Leu	Gly	Ser	Arg	Phe	Ala	Ser	Gln	Leu	Pro			
			440					445					450					
GCC	GCC	ATC	ACC	GGC	ACC	GGC	TCC	CTC	GAC	GAG	GCA	CTG	CGC	GAC	GCC			1506
Ala	Ala	Ile	Thr	Gly	Thr	Gly	Ser	Leu	Asp	Glu	Ala	Leu	Arg	Asp	Ala			
		455					460					465						
ACA	CCC	CAA	CAG	GCC	GGG	CAG	GTC	CAC	GAC	GCG	TTC	GCC	GAC	GCG	GTG			1554

03806.0054-04.seq1st

Thr	Pro	Gln	Gln	Ala	Gly	Gln	Val	His	Asp	Ala	Phe	Ala	Asp	Ala	Val		
470						475					480						
AAC	ACC	AGC	CAA	CTC	ATC	GGC	GCC	GCC	GCC	GTG	TTC	ACC	GGC	GGC	CTG	1602	
Asn	Thr	Ser	Gln	Leu	Ile	Gly	Ala	Ala	Ala	Val	Phe	Thr	Gly	Gly	Leu		
485					490					495					500		
CTC	GCC	GCG	CTG	CTG	CTG	CAC	CGC	GCC	GAC	CGC	AAG	GCC	GCC	CCC	CAG	1650	
Leu	Ala	Ala	Leu	Leu	Leu	His	Arg	Ala	Asp	Arg	Lys	Ala	Ala	Pro	Gln		
				505					510					515			
CCC	ACC	GCC	CCC	ACC	CCC	GAA	CCC	ACC	ACC	ACC	GCC	TGACCCCCGG				1696	
Pro	Thr	Ala	Pro	Thr	Pro	Glu	Pro	Thr	Thr	Thr	Ala						
			520					525									
CCCGCCGGGC	ACCACACAAC	CCACGGCCCC	ACCCCTGCGG	CTCCCCACCG	GGACCCACAG	1756											
GGGCGGGGCC	GTGCCGCTGC	CCTGCCACAC	CACACAGCCC	CCACACACAC	AGCCCCCGCA	1816											
CGGCCGACAG	CGCCGGG					1833											

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 695 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 212..695
- (D) OTHER INFORMATION: /product= "Gene Snac"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CTCGAGCCGC	GCCCCAGGT	GCTGGTGTCTG	CTCGCCGTGG	AGAAGGGCGC	CGACGGCACC	60											
GCGCCGCCGG	ACCGGCTGCT	GATCCACGAC	GGCTTCCCCT	GGGGCCGCGC	CGCCCCGCGC	120											
GAAGCGGAGC	TGCCCACCGG	GCACCGCGCC	CTGCCGGCCC	TGGCCGGCGC	CGCCCGCTGA	180											
GGCGCGGCAA	CCACCAACAG	AAGGAGCCCC	C	GTG	ACA	GGA	GCC	GAC	GAC	CCG	232						
			1	Val	Thr	Gly	Ala	Asp	Asp	Pro							
GCA	AGG	CCC	GCG	GTC	GGC	CCG	CAG	AGT	TTC	CGA	GAC	GCG	ATG	GCG	CAG	280	
Ala	Arg	Pro	Ala	Val	Gly	Pro	Gln	Ser	Phe	Arg	Asp	Ala	Met	Ala	Gln		
	10					15						20					
CTG	GCG	TCG	CCC	GTC	ACC	GTC	GTA	ACC	GTC	CTC	GAC	GCG	GCC	GGA	CGC	328	
Leu	Ala	Ser	Pro	Val	Thr	Val	Val	Thr	Val	Leu	Asp	Ala	Ala	Gly	Arg		

03806.0054-04.seq1st

25	30	35	
CGC CAC GGC TTC ACG GCC GGC TCG GTG GTC TCT GTG TCG CTG GAC CCG			376
Arg His Gly Phe Thr Ala Gly Ser Val Val Ser Val Ser Leu Asp Pro			
40 45 50 55			
CCG CTG GTG ATG GTC GGC ATC GCG CTC ACC TCC AGC TGC CAC ACG GCG			424
Pro Leu Val Met Val Gly Ile Ala Leu Thr Ser Ser Cys His Thr Ala			
60 65 70			
ATG GCC GCC GCC GCC GAG TTC TGC GTC AGC ATC CTC GGC GAG GAC CAG			472
Met Ala Ala Ala Ala Glu Phe Cys Val Ser Ile Leu Gly Glu Asp Gln			
75 80 85			
CGC GCC GTC GCG AAG CCG TGC GCG ACG CAC GGC GCC GAC CGG TTC GCG			520
Arg Ala Val Ala Lys Arg Cys Ala Thr His Gly Ala Asp Arg Phe Ala			
90 95 100			
GGC GGC GAG TTC GCC GCC TGG GAC GGT ACG GGG GTG CCC TAC CTG CCG			568
Gly Gly Glu Phe Ala Ala Trp Asp Gly Thr Gly Val Pro Tyr Leu Pro			
105 110 115			
GAC GCC AAG GTC GTC CTG CGC TGC CGC ACC ACG GAC GTG GTG CGC GCC			616
Asp Ala Lys Val Val Leu Arg Cys Arg Thr Thr Asp Val Val Arg Ala			
120 125 130 135			
GGC GAC CAC GAC CTG GTG CTC GGC ACG CCC GTG GAG ATC CGC ACG GGC			664
Gly Asp His Asp Leu Val Leu Gly Thr Pro Val Glu Ile Arg Thr Gly			
140 145 150			
GAC CCG GCG AAG CCA CCC CTG CTG TGG TAC C			695
Asp Pro Ala Lys Pro Pro Leu Leu Trp Tyr			
155 160			

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 640 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 1..640
 (D) OTHER INFORMATION: /product= "gene Snad"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

GCG ACC GCC CGG CTC ATC GGC CCG CTG CCG CGC CGG CTG GGC CTC CAG	48
Ala Thr Ala Arg Leu Ile Gly Pro Leu Pro Arg Arg Leu Gly Leu Gln	
1 5 10 15	

03806.0054-04.seq1st

GTG CAC CAG GTG ATG ACG GGC GCG TTC GCG CAG GCC CTC GCC CGC TGG Val His Gln Val Met Thr Gly Ala Phe Ala Gln Ala Leu Ala Arg Trp 20 25 30	96
CGG GGC AGC CGC GCC GTC ACC TTC GAC GTG GAG ACC CAC GGA CGG CAC Arg Gly Ser Arg Ala Val Thr Phe Asp Val Glu Thr His Gly Arg His 35 40 45	144
GGC CGC GAC GAA CTG TTC CGT ACC GTC GGC TGG TTC ACC TCC ATC CAC Gly Arg Asp Glu Leu Phe Arg Thr Val Gly Trp Phe Thr Ser Ile His 50 55 60	192
CCC GTC GTC CTG GGC GCG GAC CGC TCC GTG CAC CCC GAG CAG TAC CTC Pro Val Val Leu Gly Ala Asp Arg Ser Val His Pro Glu Gln Tyr Leu 65 70 75 80	240
GCC CAG ATC GGC GCG GCG CTG ACC GCC GTA CCG GAC GGC GGC GTC GGC Ala Gln Ile Gly Ala Ala Leu Thr Ala Ala Pro Asp Gly Gly Val Gly 85 90 95	288
TTC GGC GCC TGC CGC GAG TTC TCC CCG GAC GCC GGG CTG CGC ACT CTG Phe Gly Ala Cys Arg Glu Phe Ser Pro Asp Ala Gly Leu Arg Thr Leu 100 105 110	336
CTG CGT GAC CTG CCG CCC GCC CTG GTG TGC TTC AAC TAC TAC GGT CAG Leu Arg Asp Leu Pro Pro Ala Leu Val Cys Phe Asn Tyr Tyr Gly Gln 115 120 125	384
GCC GAC CAG TTG AGC CCG AAC GGC GGT TTC CGT ATG TCG GGC CGT CCC Ala Asp Gln Leu Ser Pro Asn Gly Gly Phe Arg Met Ser Gly Arg Pro 130 135 140	432
ATC CCG CGC GAG CAC TCC GCC CGC TGC GAG CGC GTC TAC GGC ATC GAG Ile Pro Arg Glu His Ser Ala Arg Cys Glu Arg Val Tyr Gly Ile Glu 145 150 155 160	480
GTG TAC GGC ATC GTC CAC GGC GGC CGC CTG CGC ATG GGC CTG ACC TGG Val Tyr Gly Ile Val His Gly Gly Arg Leu Arg Met Gly Leu Thr Trp 165 170 175	528
GTG CCG AGC CCG GCG GAC GGT GTG GAC GAG GCC GGC GTC GAC GCG CTC Val Pro Ser Pro Ala Asp Gly Val Asp Glu Ala Gly Val Asp Ala Leu 180 185 190	576
GTG GAG CAG ATG AGC TGG GTG CTG GCC ACG CTC GCG GGC GCC GAC CCG Val Glu Gln Met Ser Trp Val Leu Ala Thr Leu Ala Gly Ala Asp Pro 195 200 205	624
CAC GCC GTG ACC CCG G His Ala Val Thr Pro 210	640

(2) INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 645 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

03806.0054-04.seqlst

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 61..645

(D) OTHER INFORMATION: /product= "gene papA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

```

GGCGTCAAGA ACCTGCCGCT GACCGTACGG CGCGGCTGAC ACAGACAAGG GGGCCACCTG      60
GTG CGC ACC GTG CGA ACC CTG CTG ATC GAC AAC TAC GAC TCG TTC ACC      108
Val Arg Thr Val Arg Thr Leu Leu Ile Asp Asn Tyr Asp Ser Phe Thr
1      5      10      15
TAC AAC CTC TTC CAG ATG CTG GCC GAG GTG AAC GGC GCC GCT CCG CTC      156
Tyr Asn Leu Phe Gln Met Leu Ala Glu Val Asn Gly Ala Ala Pro Leu
20      25      30
GTC GTC CGC AAC GAC GAC ACC CGC ACC TGG CAG GCC CTG GCG CCG GGC      204
Val Val Arg Asn Asp Asp Thr Arg Thr Trp Gln Ala Leu Ala Pro Gly
35      40      45
GAC TTC GAC AAC GTC GTC GTC TCA CCC GGC CCC GGC CAC CCC GCC ACC      252
Asp Phe Asp Asn Val Val Val Ser Pro Gly Pro Gly His Pro Ala Thr
50      55      60
GAC ACC GAC CTG GGC CTC AGC CGC CGG GTG ATC ACC GAA TGG GAC CTG      300
Asp Thr Asp Leu Gly Leu Ser Arg Arg Val Ile Thr Glu Trp Asp Leu
65      70      75      80
CCG CTG CTC GGG GTG TGC CTG GGC CAC CAG GCC CTG TGC CTG CTC GCC      348
Pro Leu Leu Gly Val Cys Leu Gly His Gln Ala Leu Cys Leu Leu Ala
85      90      95
GGC GCC GCC GTC GTC CAC GCA CCC GAA CCC TTT CAC GGC CGC ACC AGC      396
Gly Ala Ala Val Val His Ala Pro Glu Pro Phe His Gly Arg Thr Ser
100      105      110
GAC ATC CGC CAC GAC GGG CAG GGC CTG TTC GCG AAC ATC CCC TCC CCG      444
Asp Ile Arg His Asp Gly Gln Gly Leu Phe Ala Asn Ile Pro Ser Pro
115      120      125
CTG ACC GTG GTC CGC TAC CAC TCG CTG ACC GTC CGG CAA CTG CCC GCC      492
Leu Thr Val Val Arg Tyr His Ser Leu Thr Val Arg Gln Leu Pro Ala
130      135      140
GAC CTG CGC GCC ACC GCC CAC ACC GCC GAC GGG CAG CTG ATG GCC GTC      540
Asp Leu Arg Ala Thr Ala His Thr Ala Asp Gly Gln Leu Met Ala Val
145      150      155      160
GCC CAC CGC CAC CTG CCC CGC TTC GGC GTG CAG TTC CAC CCC GAA TCG      588
Ala His Arg His Leu Pro Arg Phe Gly Val Gln Phe His Pro Glu Ser
165      170      175
ATC AGC AGC GAA CAC GGC CAC CGG ATG CTC GCC AAC TTC CGC GAC CTG      636

```

Ile Ser Ser Glu His Gly His Arg Met Leu Ala Asn Phe Arg Asp Leu
180 185 190

TCC CTG CGC
Ser Leu Arg
195

645

(2) INFORMATION FOR SEQ ID NO: 10:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1052 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 84..962
(D) OTHER INFORMATION: /product= "Gene PapM"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:

CTCGAGGACG AGTGGATCGC CTCCGGCGGC GCCCCCGTCC CCACGCCCCGT GCACGCGTCC	60
GCGTCCGCGC GGGGGGCCGT GTC GTG ACC GCC GCC GCA CCC ACC CTC GCC	110
Val Thr Ala Ala Ala Pro Thr Leu Ala	
1 5	
CAG GCG CTG GAC GAG GCC ACC GGG CAG CTG ACC GGC GCC GGG ATC ACC	158
Gln Ala Leu Asp Glu Ala Thr Gly Gln Leu Thr Gly Ala Gly Ile Thr	
10 15 20 25	
GCC GAC GCC GCC CGG GCC GAC ACC CGG CTG CTG GCC GCC CAC GCC TGC	206
Ala Asp Ala Ala Arg Ala Asp Thr Arg Leu Leu Ala Ala His Ala Cys	
30 35 40	
CAG GTC GCC CCG GGG GAC CTC GAC ACC TGC CTG GCC GGC CCG GTG CCG	254
Gln Val Ala Pro Gly Asp Leu Asp Thr Cys Leu Ala Gly Pro Val Pro	
45 50 55	
CCC CGG TTC TGG CAC TAC GTC CGG CGC CGT CTG ACC CGC GAA CCC GCC	302
Pro Arg Phe Trp His Tyr Val Arg Arg Arg Leu Thr Arg Glu Pro Ala	
60 65 70	
GAA CGC ATC GTC GGC CAC GCC TAC TTC ATG GGC CAC CGC TTC GAC CTG	350
Glu Arg Ile Val Gly His Ala Tyr Phe Met Gly His Arg Phe Asp Leu	
75 80 85	
GCC CCC GGC GTC TTC GTC CCC AAA CCC GAG ACC GAG GAG ATC ACC CGG	398
Ala Pro Gly Val Phe Val Pro Lys Pro Glu Thr Glu Glu Ile Thr Arg	
90 95 100 105	

03806.0054-04.seq1st

GAC Asp	GCC Ala	ATC Ile	GCC Ala	CGC Arg 110	CTG Leu	GAG Glu	GCC Ala	CTC Leu	GTC Val 115	CGC Arg	CGC Arg	GGC Gly	ACC Thr	ACC Thr 120	GCA Ala	446
CCC Pro	CTG Leu	GTC Val	GTC Val 125	GAC Asp	CTG Leu	TGC Cys	GCC Ala	GGA Gly 130	CCG Pro	GGC Gly	ACC Thr	ATG Met	GCC Ala 135	GTC Val	ACC Thr	494
CTG Leu	GCC Ala	CGC Arg 140	CAC His	GTA Val	CCG Pro	GCC Ala	GCC Ala 145	CGC Arg	GTC Val	CTG Leu	GGC Gly	ATC Ile 150	GAA Glu	CTC Leu	TCC Ser	542
CAG Gln 155	GCC Ala	GCC Ala	GCC Ala	CGC Arg	GCC Ala	GCC Ala 160	CGG Arg	CGC Arg	AAC Asn	GCC Ala	CGC Arg 165	GGC Gly	ACC Thr	GGC Gly	GCC Ala	590
CGC Arg 170	ATC Ile	GTG Val	CAG Gln	GGC Gly	GAC Asp 175	GCC Ala	CGC Arg	GAC Asp	GCC Ala	TTC Phe 180	CCC Pro	GAA Glu	CTG Leu	AGC Ser	GGC Gly 185	638
ACC Thr	GTC Val	GAC Asp	CTC Leu	GTC Val 190	GTC Val	ACC Thr	AAC Asn	CCG Pro	CCC Pro 195	TAC Tyr	ATC Ile	CCC Pro	ATC Ile	GGA Gly 200	CTG Leu	686
CGC Arg	ACC Thr	TCC Ser	GCA Ala 205	CCC Pro	GAA Glu	GTG Val	CTC Leu	GAG Glu 210	CAC His	GAC Asp	CCG Pro	CCG Pro	CTG Leu 215	GCC Ala	CTG Leu	734
TGG Trp	GCC Ala	GGG Gly 220	GAG Glu	GAG Glu	GGC Gly	CTC Leu	GGC Gly 225	ATG Met	ATC Ile	CGC Arg	GCC Ala	ATG Met 230	GAA Glu	CGC Arg	ACC Thr	782
GCG Ala 235	GCC Ala	CGG Arg	CTG Leu	CTG Leu	GCC Ala	CCC Pro 240	GGC Gly	GGC Gly	GTC Val	CTG Leu	CTC Leu 245	CTC Leu	GAA Glu	CAC His	GGC Gly	830
TCC Ser 250	TAC Tyr	CAA Gln	CTC Leu	GCC Ala	TCC Ser 255	GTG Val	CCC Pro	GCC Ala	CTG Leu	TTC Phe 260	CGC Arg	GCA Ala	ACC Thr	GGC Gly	CGC Arg 265	878
TGG Trp	AGC Ser	CAC His	GCC Ala	TCG Ser 270	TCC Ser	CGT Arg	CCC Pro	ACC Thr	TGC Cys 275	AAC Asn	GAC Asp	GGC Gly	TGC Cys	CTG Leu 280	ACC Thr	926
GCC Ala	GTA Val	CGC Arg	AAC Asn 285	CAC His	ACC Thr	TGC Cys	GCA Ala	CCG Pro 290	CCC Pro	GCC Ala	TGACACGGCG TCACGGCACG					979
GCCGGCCTGT CGGCAACGAC CCTACGCCAT TGACAAACCG ACCGTGCCGT TTTTTTAATG																1039
TCGGGGTGGC GGA																1052

(2) INFORMATION FOR SEQ ID NO: 11:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 227 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: double
 - (D) TOPOLOGY: linear

03806.0054-04.seqlst

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 3..227

(D) OTHER INFORMATION: /product= "Partie du gene Snbc"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

AG ATC TTC GAG CAC AAG ACC GTC GCC CAG CTC GCA CCC GTC GCC GAG	47
Ile Phe Glu His Lys Thr Val Ala Gln Leu Ala Pro Val Ala Glu	
1 5 10 15	
ACG CTC GCC GAC ACC ACC CGC GAG GAA CCC GCC GCC GTC GCC GCG ACC	95
Thr Leu Ala Asp Thr Thr Arg Glu Glu Pro Ala Ala Val Ala Ala Thr	
20 25 30	
GGC GAC GTA CCG CTC ACC CCG ATC ATG CAC TGG CTG CGC GAA CGC GGC	143
Gly Asp Val Pro Leu Thr Pro Ile Met His Trp Leu Arg Glu Arg Gly	
35 40 45	
GGC CCC GTC GAC GCG TTC AGC CAG ACG ATG GCC GTC ACC GTC CCC GCC	191
Gly Pro Val Asp Ala Phe Ser Gln Thr Met Ala Val Thr Val Pro Ala	
50 55 60	
GGC CTG GAC CGG GAA CGG CTC GTG GCC GCC CTG CAG	227
Gly Leu Asp Arg Glu Arg Leu Val Ala Ala Leu Gln	
65 70 75	

(2) INFORMATION FOR SEQ ID NO: 12:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 247 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: double

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..247

(D) OTHER INFORMATION: /product= "Partie du gene Snbc"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 12:

03806.0054-04.seq1st

CTC GAG TAC GAC ACC GCC CTG TAC GAG CGG GCC ACC GCC GAA GCC CTC	48
Leu Glu Tyr Asp Thr Ala Leu Tyr Glu Arg Ala Thr Ala Glu Ala Leu	
1 5 10 15	
ACC GGC CGG CTG CTG CGG CTC CTC GAC GCC GTC GTC ACC GAC CCG CAG	96
Thr Gly Arg Leu Leu Arg Leu Leu Asp Ala Val Val Thr Asp Pro Gln	
20 25 30	
GCG CCG GTC GGC TCC CAC GAC CTC CTC GAA GAG GCC GAA CAC GCC CGC	144
Ala Pro Val Gly Ser His Asp Leu Leu Glu Glu Ala Glu His Ala Arg	
35 40 45	
CTG GCA GCC TTC AAC GAC ACC GCC CGG CCC GTG CCG CGA GCC GGC CTC	192
Leu Ala Ala Phe Asn Asp Thr Ala Arg Pro Val Pro Arg Ala Gly Leu	
50 55 60	
GCC GAA CTC TTC ACC GCC CAG GCC CGC CGC ACC GCC GAT GCG GTC GCC	240
Ala Glu Leu Phe Thr Ala Gln Ala Arg Arg Thr Ala Asp Ala Val Ala	
65 70 75 80	
GTC GTC G	247
Val Val	

(2) INFORMATION FOR SEQ ID NO: 13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 192 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

- (A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 3..192
- (D) OTHER INFORMATION: /product= "Partie du gene SnbD"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 13:

GC ATG CCC CCC GTC ACC CCC TAC CGC GCC TAC CTG GCC CAC CTC GCC	47
Met Pro Pro Val Thr Pro Tyr Arg Ala Tyr Leu Ala His Leu Ala	
1 5 10 15	
GGC CGT GAC GAC GAC GCC GCC CGC GCC GCG TGG CGG ACC GCC CTC GCG	95
Gly Arg Asp Asp Asp Ala Ala Arg Ala Ala Trp Arg Thr Ala Leu Ala	
20 25 30	
GAC CTG GAG GAG CCG AGC CTC GTC GCG GGC GCC GGA GCA GGC CGC GGC	143
Asp Leu Glu Glu Pro Ser Leu Val Ala Gly Ala Gly Ala Gly Arg Gly	
35 40 45	

03806.0054-04.seq1st

GCC	GCC	GAC	GGC	TCC	GCC	CTG	CCC	GGC	CAG	ATC	CCC	GGT	TAC	CGA	GCT	C	192
Ala	Ala	Asp	Gly	Ser	Ala	Leu	Pro	Gly	Gln	Ile	Pro	Gly	Tyr	Arg	Ala		
		50					55					60					

(2) INFORMATION FOR SEQ ID NO: 14:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 474 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

(A) NAME/KEY: CDS

(B) LOCATION: 1..474

(D) OTHER INFORMATION: /product= "Partie du gene snbD"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 14:

CTG	CAG	GTC	GAG	GGC	CGG	CCC	GCG	CAC	CTG	GAA	CTG	CCC	TGC	GAC	CAC	48
Leu	Gln	Val	Glu	Gly	Arg	Pro	Ala	His	Leu	Glu	Leu	Pro	Cys	Asp	His	
1				5					10					15		
CCC	CGG	CCC	GCC	GTC	GCC	ACC	CAC	CGC	GGC	GCC	ACC	GTG	CCC	TTC	CAC	96
Pro	Arg	Pro	Ala	Val	Ala	Thr	His	Arg	Gly	Ala	Thr	Val	Pro	Phe	His	
			20					25					30			
ATC	GAC	GCC	GGC	CTC	CAC	GAG	AAG	CTG	ACC	GCG	CTC	TCC	AAG	GCC	TGC	144
Ile	Asp	Ala	Gly	Leu	His	Glu	Lys	Leu	Thr	Ala	Leu	Ser	Lys	Ala	Cys	
		35					40					45				
GAC	AGC	AGC	CTG	TTC	ATG	GTG	CTC	CAG	GCC	GCG	GTC	GCC	GCC	CTG	CTC	192
Asp	Ser	Ser	Leu	Phe	Met	Val	Leu	Gln	Ala	Ala	Val	Ala	Ala	Leu	Leu	
		50				55					60					
ACC	CGG	CAC	GGC	GCC	GGC	ACC	GAC	ATC	CCC	GTC	GGC	AGC	CCC	GTC	GCC	240
Thr	Arg	His	Gly	Ala	Gly	Thr	Asp	Ile	Pro	Val	Gly	Ser	Pro	Val	Ala	
		65			70				75						80	
GGC	CGC	ACC	GAC	GAC	GCC	CTC	GAC	GAC	CTG	GTG	GGC	TTC	TTC	GTC	AAC	288
Gly	Arg	Thr	Asp	Asp	Ala	Leu	Asp	Asp	Leu	Val	Gly	Phe	Phe	Val	Asn	
			85						90					95		
ACC	CTC	GTC	CTG	CGC	ACC	GAC	ACC	TCC	GGC	GAC	CCC	ACC	TTC	CGC	GAA	336
Thr	Leu	Val	Leu	Arg	Thr	Asp	Thr	Ser	Gly	Asp	Pro	Thr	Phe	Arg	Glu	
			100					105					110			
CTC	GTC	GCA	CGC	GTG	CGG	CAG	TTC	GAC	CTC	GCC	GCC	TAC	ACG	CAC	CAG	384
Leu	Val	Ala	Arg	Val	Arg	Gln	Phe	Asp	Leu	Ala	Ala	Tyr	Thr	His	Gln	
		115				120						125				
GAC	ATG	CCG	TTC	GAA	AAG	CTC	GTC	GAA	GAG	GTC	AAC	CCC	GAG	CGC	TCC	432

03806.0054-04.seqlst

Asp	Met	Pro	Phe	Glu	Lys	Leu	Val	Glu	Glu	Val	Asn	Pro	Glu	Arg	Ser
130						135					140				

CTG	GCC	CGC	AAC	CCG	CTC	TTC	CAG	GTC	GTC	CTG	GCG	CTG	CAG		474
Leu	Ala	Arg	Asn	Pro	Leu	Phe	Gln	Val	Val	Leu	Ala	Leu	Gln		
145					150					155					

(2) INFORMATION FOR SEQ ID NO: 15:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 485 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: double
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:
 (A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:
 (A) NAME/KEY: CDS
 (B) LOCATION: 3..485
 (D) OTHER INFORMATION: /product= "Partie du gene SnbE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 15:

GC	ATG	CCG	CGC	TCC	CTC	GAC	CTG	TAC	GTC	GCA	CTG	CTC	GCC	GTC	CTC	47
Met	Pro	Arg	Ser	Leu	Asp	Leu	Tyr	Val	Ala	Leu	Leu	Ala	Val	Leu		
1				5					10					15		
AAG	ACC	GGC	GCC	GCC	TAC	CTG	CCC	GTC	GAC	ATC	TCC	TAC	CCG	GCC	GAA	95
Lys	Thr	Gly	Ala	Ala	Tyr	Leu	Pro	Val	Asp	Ile	Ser	Tyr	Pro	Ala	Glu	
			20						25					30		
CGC	ATC	GCG	TTC	ATG	ATC	GAG	GAC	GCC	CGC	CCG	GTG	ACC	GTC	CTC	GAC	143
Arg	Ile	Ala	Phe	Met	Ile	Glu	Asp	Ala	Arg	Pro	Val	Thr	Val	Leu	Asp	
			35					40					45			
CGC	CTG	CCC	GAC	GAC	CTG	GGC	GCC	TAC	CGG	GAC	ACC	GAC	CTC	ACC	GAC	191
Arg	Leu	Pro	Asp	Asp	Leu	Gly	Ala	Tyr	Arg	Asp	Thr	Asp	Leu	Thr	Asp	
		50				55						60				
GCC	GAC	CGC	ACG	GCG	CCG	CTA	CGG	CCC	GAA	CAC	CCG	GCG	TAC	GTC	ATC	239
Ala	Asp	Arg	Thr	Ala	Pro	Leu	Arg	Pro	Glu	His	Pro	Ala	Tyr	Val	Ile	
	65					70					75					
CAC	ACC	TCC	GGC	TCC	ACC	GGC	ACC	CCC	AAG	GCC	GTC	GTC	ATG	CCC	CAC	287
His	Thr	Ser	Gly	Ser	Thr	Gly	Thr	Pro	Lys	Ala	Val	Val	Met	Pro	His	
	80				85					90					95	
GCC	GGC	CTG	GTC	AAC	CTG	CTG	ACC	TGG	CAC	GCC	CGC	CGC	TTC	CCC	GGC	335
Ala	Gly	Leu	Val	Asn	Leu	Leu	Thr	Trp	His	Ala	Arg	Arg	Phe	Pro	Gly	
				100					105					110		
GGC	ACC	GGG	GTG	CGC	ACC	GCC	CAG	TTC	ACC	GCC	ATC	GGC	TTC	GAC	TTC	383

03806.0054-04.seq1st

Gly Thr Gly Val Arg Thr Ala Gln	Phe Thr Ala Ile Gly Phe Asp Phe	
115	120 125	
TCG GTG CAG GAG ATC CTC TCC CCG CTC GTC ATG GGC AAG ACC CTC GCC		431
Ser Val Gln Glu Ile Leu Ser Pro Leu Val Met Gly Lys Thr Leu Ala		
130	135 140	
GTG CCC TCG GAA GAG GTC CGC CAC AGC GCC GAA CTG CTG GCC GGC TGG		479
Val Pro Ser Glu Glu Val Arg His Ser Ala Glu Leu Leu Ala Gly Trp		
145	150 155	
CTC GAG		485
Leu Glu		
160		

(2) INFORMATION FOR SEQ ID NO: 16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 291 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(iii) HYPOTHETICAL: NO

(iv) ANTI-SENSE: NO

(vi) ORIGINAL SOURCE:

(A) ORGANISM: *S.pristinaespiralis*

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 1..291
- (D) OTHER INFORMATION: /product= "Partie du gene SnbE"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 16:

CTG CAG GCC GAG GGC GCC GAA GTG AGC CTG CTG GCC GTC CTC GAC GGC	48
Leu Gln Ala Glu Gly Ala Glu Val Ser Leu Leu Ala Val Leu Asp Gly	
1 5 10 15	
TAC CCC GAC GCC TAC GAC GGC ACC GAG CAC GAG GTC GGC GAG GAA CAG	96
Tyr Pro Asp Ala Tyr Asp Gly Thr Glu His Glu Val Gly Glu Glu Gln	
20 25 30	
GTC CTG GCG ATC CTC CTC AAC GCC GCC GGC GTC GAC CGG GCC CAG GCC	144
Val Leu Ala Ile Leu Leu Asn Ala Ala Gly Val Asp Arg Ala Gln Ala	
35 40 45	
TTC GGC GAC GCC CCC CTC CAA CGG GCC GCC GTG CTC GAG AAG CTG CGC	192
Phe Gly Asp Ala Pro Leu Gln Arg Ala Ala Val Leu Glu Lys Leu Arg	
50 55 60	
GAC AGC GGC AGC GCC CTG GGC AAC CTC GAC GAC GAC GCG GTC GGC CGC	240
Asp Ser Gly Ser Ala Leu Gly Asn Leu Asp Asp Asp Ala Val Gly Arg	
65 70 75 80	
ATG GTC ACC GTC TTC CTC AAC AAC ACG CGC CTC ATC CAG AAC TTC CGG	288

03806.0054-04.seq1st

291

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 422 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

Met 1	Thr	Ala	Pro	Arg 5	Arg	Arg	Ile	Thr	Leu 10	Ala	Gly	Ile	Ile	Asp 15	Gly
Pro	Gly	Gly	His 20	Val	Ala	Ala	Trp	Arg 25	His	Pro	Ala	Thr	Lys 30	Ala	Asp
Ala	Gln	Leu 35	Asp	Phe	Glu	Phe	His 40	Arg	Asp	Asn	Ala	Arg 45	Thr	Leu	Glu
Arg	Gly 50	Leu	Phe	Asp	Ala	Val 55	Phe	Ile	Ala	Asp	Ile 60	Val	Ala	Val	Trp
Gly 65	Thr	Arg	Leu	Asp	Ser 70	Leu	Cys	Arg	Thr	Ser 75	Arg	Thr	Glu	His	Phe 80
Glu	Pro	Leu	Thr	Leu 85	Leu	Ala	Ala	Tyr	Ala 90	Ala	Val	Thr	Glu	His 95	Ile
Gly	Leu	Cys	Ala 100	Thr	Ala	Thr	Thr	Thr 105	Tyr	Asn	Glu	Pro	Ala 110	His	Ile
Ala	Ala	Arg 115	Phe	Ala	Ser	Leu	Asp 120	His	Leu	Ser	Gly	Gly 125	Arg	Ala	Gly
Trp	Asn 130	Val	Val	Thr	Ser	Ala 135	Ala	Pro	Trp	Glu	Ser 140	Ala	Asn	Phe	Gly
Phe 145	Pro	Glu	His	Leu	Glu 150	His	Gly	Lys	Arg	Tyr 155	Glu	Arg	Ala	Glu	Glu 160
Phe	Ile	Asp	Val	Val 165	Lys	Lys	Leu	Trp	Asp 170	Ser	Asp	Gly	Arg	Pro 175	Val
Asp	His	Arg	Gly 180	Thr	His	Phe	Glu	Ala 185	Pro	Gly	Pro	Leu	Gly 190	Ile	Ala
Arg	Pro	Pro 195	Gln	Gly	Arg	Pro	Val 200	Ile	Ile	Gln	Ala	Gly 205	Ser	Ser	Pro
Val	Gly 210	Arg	Glu	Phe	Ala	Ala 215	Arg	His	Ala	Glu	Val 220	Ile	Phe	Thr	Arg
His 225	Asn	Arg	Leu	Ser	Asp 230	Ala	Gln	Asp	Phe	Tyr 235	Gly	Asp	Leu	Lys	Ala 240
Arg	Val	Ala	Arg	His	Gly	Arg	Asp	Pro	Glu	Lys	Val	Leu	Val	Trp	Pro

03806.0054-04.seq1st

245 250 255
 Thr Leu Ala Pro Ile Val Ala Ala Thr Asp Thr Glu Ala Lys Gln Arg
 260 265 270
 Leu Gln Glu Leu Gln Asp Leu Thr His Asp His Val Ala Leu Arg Thr
 275 280 285
 Leu Gln Asp His Leu Gly Asp Val Asp Leu Ser Ala Tyr Pro Ile Asp
 290 295 300
 Gly Pro Val Pro Asp Ile Pro Tyr Thr Asn Gln Ser Gln Ser Thr Thr
 305 310 315 320
 Glu Arg Leu Ile Gly Leu Ala Arg Arg Glu Asn Leu Ser Ile Arg Glu
 325 330 335
 Leu Ala Leu Arg Leu Met Gly Asp Ile Val Val Gly Thr Pro Glu Gln
 340 345 350
 Leu Ala Asp His Met Glu Ser Trp Phe Thr Gly Arg Gly Ala Asp Gly
 355 360 365
 Phe Asn Ile Asp Phe Pro Tyr Leu Pro Gly Ser Ala Asp Asp Phe Val
 370 375 380
 Asp His Val Val Pro Glu Leu Gln Arg Arg Gly Leu Tyr Arg Ser Gly
 385 390 395 400
 Tyr Glu Gly Thr Thr Leu Arg Ala Asn Leu Gly Ile Asp Ala Pro Arg
 405 410 415
 Lys Ala Gly Ala Ala Ala
 420

(2) INFORMATION FOR SEQ ID NO: 18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 277 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 18:

Met Thr Ala Pro Ile Leu Val Ala Thr Leu Asp Thr Arg Gly Pro Ala
 1 5 10 15
 Ala Thr Leu Gly Thr Ile Thr Arg Ala Val Arg Ala Ala Glu Ala Ala
 20 25 30
 Gly Phe Asp Ala Val Leu Ile Asp Asp Arg Ala Ala Ala Gly Val Gln
 35 40 45
 Gly Arg Phe Glu Thr Thr Thr Leu Thr Ala Ala Leu Ala Ala Val Thr
 50 55 60
 Glu His Ile Gly Leu Ile Thr Ala Pro Leu Pro Ala Asp Gln Ala Pro
 65 70 75 80
 Tyr His Val Ser Arg Ile Thr Ala Ser Leu Asp His Leu Ala His Gly
 85 90 95

03806.0054-04.seq1st

Arg Thr Gly Trp Leu Ala Ser Thr Asp Thr Thr Asp Pro Glu Gly Arg
100 105 110
Thr Gly Glu Leu Ile Asp Val Val Arg Gly Leu Trp Asp Ser Phe Asp
115 120 125
Asp Asp Ala Phe Val His Asp Arg Ala Asp Gly Leu Tyr Trp Arg Leu
130 135 140
Pro Ala Val His Gln Leu Asp His Gln Gly Arg His Phe Asp Val Ala
145 150 155 160
Gly Pro Leu Asn Val Ala Arg Pro Pro Gln Gly His Pro Val Val Ala
165 170 175
Val Thr Gly Pro Ala Leu Ala Ala Ala Asp Leu Val Leu Leu Asp
180 185 190
Glu Ala Ala Asp Ala Ala Ser Val Lys Gln Gln Ala Pro His Ala Lys
195 200 205
Ile Leu Leu Pro Leu Pro Gly Pro Ala Ala Glu Leu Pro Ala Asp Ser
210 215 220
Pro Ala Asp Gly Phe Thr Val Ala Leu Thr Gly Ser Asp Asp Pro Val
225 230 235 240
Leu Ala Ala Leu Ala Ala Arg Pro Gly Arg Pro Asp Arg Thr Ala Ala
245 250 255
Thr Thr Leu Arg Glu Arg Leu Gly Leu Ala Arg Pro Glu Ser Arg His
260 265 270
Ala Leu Thr Thr Ala
275

(2) INFORMATION FOR SEQ ID NO: 19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 402 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 19:

Met Ser Arg Arg Leu Phe Thr Ser Glu Ser Val Thr Glu Gly His Pro
1 5 10 15
Asp Lys Ile Ala Asp Gln Ile Ser Asp Thr Val Leu Asp Ala Leu Leu
20 25 30
Arg Glu Asp Pro Ala Ser Arg Val Ala Val Glu Thr Leu Ile Thr Thr
35 40 45
Gly Gln Val His Ile Ala Gly Glu Val Thr Thr Lys Ala Tyr Ala Pro
50 55 60
Ile Ala Gln Leu Val Arg Asp Thr Ile Leu Ala Ile Gly Tyr Asp Ser
65 70 75 80

03806.0054-04.seq1st

Ser Ala Lys Gly Phe Asp Gly Ala Ser Cys Gly Val Ser Val Ser Ile
 85 90 95
 Gly Ala Gln Ser Pro Asp Ile Ala Gln Gly Val Asp Ser Ala Tyr Glu
 100 105 110
 Thr Arg Val Glu Gly Glu Asp Asp Glu Leu Asp Gln Gln Gly Ala Gly
 115 120 125
 Asp Gln Gly Leu Met Phe Gly Tyr Ala Thr Asp Glu Thr Pro Ser Leu
 130 135 140
 Met Pro Leu Pro Ile Glu Leu Ala His Arg Leu Ser Arg Arg Leu Thr
 145 150 155 160
 Glu Val Arg Lys Asp Gly Thr Val Pro Tyr Leu Arg Pro Asp Gly Lys
 165 170 175
 Thr Gln Val Thr Ile Glu Tyr Gln Gly Ser Arg Pro Val Arg Leu Asp
 180 185 190
 Thr Val Val Val Ser Ser Gln His Ala Ala Asp Ile Asp Leu Gly Ser
 195 200 205
 Leu Leu Thr Pro Asp Ile Arg Glu His Val Val Glu His Val Leu Ala
 210 215 220
 Ala Leu Ala Glu Asp Gly Ile Lys Leu Glu Thr Asp Asn Tyr Arg Leu
 225 230 235 240
 Leu Val Asn Pro Thr Gly Arg Phe Glu Ile Gly Gly Pro Met Gly Asp
 245 250 255
 Ala Gly Leu Thr Gly Arg Lys Ile Ile Ile Asp Thr Tyr Gly Gly Met
 260 265 270
 Ala Arg His Gly Gly Gly Ala Phe Ser Gly Lys Asp Pro Ser Lys Val
 275 280 285
 Asp Arg Ser Ala Ala Tyr Ala Met Arg Trp Val Ala Lys Asn Val Val
 290 295 300
 Ala Ala Gly Leu Ala Ser Arg Cys Glu Val Gln Val Ala Tyr Ala Ile
 305 310 315 320
 Gly Lys Ala Glu Pro Val Gly Leu Phe Val Glu Thr Phe Gly Thr Gly
 325 330 335
 Thr Val Ala Gln Glu Arg Ile Glu Lys Ala Ile Thr Glu Val Phe Asp
 340 345 350
 Leu Arg Pro Ala Ala Ile Ile Arg Asp Leu Asp Leu Leu Arg Pro Ile
 355 360 365
 Tyr Ala Ala Thr Ala Ala Tyr Gly His Phe Gly Arg Glu Leu Pro Asp
 370 375 380
 Phe Thr Trp Glu Arg Thr Asp Arg Ala His Arg Leu Lys Ala Ala Ala
 385 390 395 400
 Gly Leu

(2) INFORMATION FOR SEQ ID NO: 20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 582 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 20:

```

Met Leu
 1
Asp Gly Cys Val Pro Trp Pro Glu Asp Val Ala Ala Lys Tyr Arg Ala
   5          10          15
Ala Gly Tyr Trp Arg Gly Glu Pro Leu Gly Met Leu Leu Gly Arg Trp
  20          25          30
Ala Glu Gln Tyr Gly Glu Arg Glu Ala Leu Val Gly Ala Asp Gly Cys
  35          40          45          50
Ser Arg Val Thr Tyr Arg Ala Leu Asp Arg Trp Cys Asp Arg Leu Ala
          55          60          65
Ala Gly Phe Ala Ala Arg Gly Ile Gly Ala Gly Glu Arg Val Leu Val
          70          75          80
Gln Leu Pro Asn Thr Pro Glu Phe Val Ala Val Cys Phe Ala Leu Phe
          85          90          95
Arg Leu Gly Ala Leu Pro Val Phe Ala Leu Pro Ala His Arg Ala Ala
 100          105          110
Glu Val Gly His Leu Leu Glu Leu Ser Gly Ala Val Ala His Ile Leu
 115          120          125          130
Pro Gly Thr Gly Thr Gly Tyr Asp His Val Ala Ala Ala Val Glu Ala
          135          140          145
Arg Ala Arg Arg Ala Arg Pro Val Gln Val Phe Val Ala Gly Glu Ala
          150          155          160
Pro Ala Val Leu Pro Glu Gly Phe Thr Ala Leu Ala Asp Val Asp Gly
          165          170          175
Asp Pro Val Ala Pro Ala Asp Val Asp Ala Phe Arg Arg Gly Val Phe
          180          185          190
Leu Leu Ser Gly Gly Thr Thr Ala Leu Pro Lys Leu Ile Pro Arg Thr
          195          200          205          210
His Asp Asp Tyr Ala Tyr Gln Cys Arg Val Thr Ala Gly Ile Cys Gly
          215          220          225
Leu Asp Ala Asp Ser Val Tyr Leu Ala Val Leu Pro Ala Glu Phe Asn
          230          235          240
Phe Pro Phe Gly Cys Pro Gly Ile Leu Gly Thr Leu His Ala Gly Gly
          245          250          255

```

03806.0054-04.seq1st

Arg Val Val Phe Ala Leu Ser Pro Gln Pro Glu Glu Cys Phe Ala Leu
260 265 270

Ile Glu Arg Glu His Val Thr Phe Thr Ser Val Ile Pro Thr Ile Val
275 280 285 290

His Leu Trp Leu Ala Ala Ala Ala Gln Gly His Gly Arg Asp Leu Gly
295 300 305

Ser Leu Gln Leu Leu Gln Val Gly Ser Ala Lys Leu His Glu Glu Leu
310 315 320

Ala Ala Arg Ile Gly Pro Glu Leu Gly Val Arg Leu Gln Gln Val Phe
325 330 335

Gly Met Ala Glu Gly Leu Leu Thr Phe Thr Arg Asp Asp Asp Pro Ala
340 345 350

Asp Val Val Leu Arg Thr Gln Gly Arg Pro Val Ser Glu Ala Asp Glu
355 360 365 370

Ile Arg Val Ala Asp Pro Asp Gly Arg Pro Val Pro Arg Gly Glu Thr
375 380 385

Gly Glu Leu Leu Thr Arg Gly Pro Tyr Thr Leu Arg Gly Tyr Tyr Arg
390 395 400

Ala Pro Glu His Asn Ala Arg Ala Phe Thr Glu Asp Gly Phe Tyr Arg
405 410 415

Ser Gly Asp Leu Val Arg Leu Thr Ala Asp Gly Gln Leu Val Val Glu
420 425 430

Gly Arg Ile Lys Asp Val Val Ile Arg Gly Gly Asp Lys Val Ser Ala
435 440 445 450

Thr Glu Val Glu Gly His Leu Gly Ala His Pro Asp Val Gln Gln Ala
455 460 465

Ala Val Val Ala Met Pro Asp Pro Val Trp Gly Glu Lys Val Cys Ala
470 475 480

Tyr Ile Val Pro Ala Pro Gly Arg Pro Ala Pro Pro Met Ala Ala Leu
485 490 495

Arg Arg Leu Leu Arg Ala Arg Gly Leu Ala Asp Tyr Lys Leu Pro Asp
500 505 510

Arg Val Glu Val Val Asp Ala Phe Pro Leu Thr Gly Leu Asn Lys Val
515 520 525 530

Asp Lys Lys Ala Leu Ala Ala Asp Ile Ala Ala Lys Thr Ala Pro Thr
535 540 545

Arg Pro Thr Thr Ala Gly His Gly Pro Thr Thr Asp Gly Asp Thr Ala
550 555 560

Gly Gly Gly Gly Ser Ala Gly Gly Val Thr Ala Ala Gly Gly Gly Arg
565 570 575

Glu Glu Ala Ala
580

(2) INFORMATION FOR SEQ ID NO: 21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 528 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 21:

```

Met Arg Thr Ser
1
Arg Ser His Asp Gln Arg Ala Pro Thr Pro Trp Arg His Pro Leu His
5      10      15      20
Ser Thr Arg Pro Ala Pro Ala Ala Asp Arg Asp Pro Arg Arg Trp Val
25      30      35
Ile Leu Gly Val Ile Cys Leu Ala Gln Leu Val Val Leu Leu Asp Asn
40      45      50
Thr Val Leu Asn Val Ala Ile Pro Val Leu Thr Thr Asp Leu Gly Ala
55      60      65
Ser Thr Ala Asp Ile Gln Trp Met Ile Asn Ala Tyr Ala Leu Val Gln
70      75      80
Ser Gly Leu Leu Leu Thr Ala Gly Ser Leu Ala Asp Arg Tyr Gly Arg
85      90      95      100
Lys Arg Leu Leu Met Leu Gly Leu Val Leu Phe Gly Ala Gly Ser Ala
105     110     115
Trp Ala Ala Phe Ala Gln Asp Ser Ala Gln Leu Ile Ala Ala Arg Ala
120     125     130
Gly Met Gly Val Gly Gly Ala Leu Leu Ala Thr Thr Thr Leu Ala Val
135     140     145
Ile Met Gln Val Phe Asp Asp Asp Glu Arg Pro Arg Ala Ile Gly Leu
150     155     160
Trp Gly Ala Ala Ser Ser Leu Gly Phe Ala Ala Gly Pro Leu Leu Gly
165     170     175     180
Gly Ala Leu Leu Asp His Phe Trp Trp Gly Ser Ile Phe Leu Ile Asn
185     190     195
Leu Pro Val Ala Leu Leu Gly Leu Leu Ala Val Ala Arg Leu Val Pro
200     205     210
Glu Thr Lys Asn Pro Glu Gly Arg Arg Pro Asp Leu Leu Gly Ala Val
215     220     225
Leu Ser Thr Leu Gly Met Val Gly Val Val Tyr Ala Ile Ile Ser Gly
230     235     240
Pro Glu His Gly Trp Thr Ala Pro Gln Val Leu Leu Pro Ala Ala Val
245     250     255     260
Ala Ala Ala Ala Leu Thr Ala Phe Val Arg Trp Glu Leu His Thr Pro

```

03806.0054-04.seq1st

265		270		275
His Pro Met	Leu Asp Met Gly Phe	Phe Thr Asp Arg Arg	Phe Asn Gly	
280		285	290	
Pro Ser Pro	Ala Glu Cys Ser	Ser Phe Gly Met Ala	Gly Ser Leu Phe	
295		300	305	
Leu Leu Thr	Gln His Leu	Gln Leu Val Leu	Gly Tyr Asp Ala Leu Gln	
310		315	320	
Ala Gly Leu Arg Thr	Ala Pro Leu Ala Leu	Thr Ile Val Ala Leu Asn		
325		330	335	340
Leu Ala Gly Leu	Gly Ala Lys Leu Leu	Ala Ala Leu Gly Thr	Ala Arg	
	345	350	355	
Ser Ile Ala	Leu Gly Met Thr Leu	Leu Ala Ala Gly Leu	Ser Ala Val	
	360	365	370	
Ala Val Gly Gly Ser Gly Pro	Asp Ala Gly Tyr Gly	Gly Met Leu Ala		
	375	380	385	
Gly Leu Leu Met Gly	Ala Gly Ile Ala Leu	Ala Met Pro Ala Met		
	390	395	400	
Ala Thr Ala Val Met	Ser Ser Ile Pro Pro	Ala Lys Ala Gly Ala Gly		
405		410	415	420
Ala Gly Val Gln Gly Thr Leu Thr	Glu Phe Gly Gly Gly Leu	Gly Val		
	425	430	435	
Ala Ile Leu Gly Ala Val Leu Gly	Ser Arg Phe Ala Ser	Gln Leu Pro		
	440	445	450	
Ala Ala Ile Thr Gly Thr Gly	Ser Leu Asp Glu Ala Leu Arg Asp Ala			
	455	460	465	
Thr Pro Gln Gln Ala Gly Gln Val His Asp Ala	Phe Ala Asp Ala Val			
	470	475	480	
Asn Thr Ser Gln Leu Ile Gly Ala Ala Ala	Val Phe Thr Gly Gly Leu			
485		490	495	500
Leu Ala Ala Leu Leu Leu His Arg Ala	Asp Arg Lys Ala Ala Pro Gln			
	505	510	515	
Pro Thr Ala Pro Thr Pro Glu Pro Thr Thr Thr Ala				
	520	525		

(2) INFORMATION FOR SEQ ID NO: 22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 161 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 22:

Val Thr Gly Ala Asp Asp Pro
1 5

03806.0054-04.seq1st

Ala Arg Pro Ala Val Gly Pro Gln Ser Phe Arg Asp Ala Met Ala Gln
 10 15 20
 Leu Ala Ser Pro Val Thr Val Val Thr Val Leu Asp Ala Ala Gly Arg
 25 30 35
 Arg His Gly Phe Thr Ala Gly Ser Val Val Ser Val Ser Leu Asp Pro
 40 45 50 55
 Pro Leu Val Met Val Gly Ile Ala Leu Thr Ser Ser Cys His Thr Ala
 60 65 70
 Met Ala Ala Ala Ala Glu Phe Cys Val Ser Ile Leu Gly Glu Asp Gln
 75 80 85
 Arg Ala Val Ala Lys Arg Cys Ala Thr His Gly Ala Asp Arg Phe Ala
 90 95 100
 Gly Gly Glu Phe Ala Ala Trp Asp Gly Thr Gly Val Pro Tyr Leu Pro
 105 110 115
 Asp Ala Lys Val Val Leu Arg Cys Arg Thr Thr Asp Val Val Arg Ala
 120 125 130 135
 Gly Asp His Asp Leu Val Leu Gly Thr Pro Val Glu Ile Arg Thr Gly
 140 145 150
 Asp Pro Ala Lys Pro Pro Leu Leu Trp Tyr
 155 160

(2) INFORMATION FOR SEQ ID NO: 23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 213 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 23:

Ala Thr Ala Arg Leu Ile Gly Pro Leu Pro Arg Arg Leu Gly Leu Gln
 1 5 10 15
 Val His Gln Val Met Thr Gly Ala Phe Ala Gln Ala Leu Ala Arg Trp
 20 25 30
 Arg Gly Ser Arg Ala Val Thr Phe Asp Val Glu Thr His Gly Arg His
 35 40 45
 Gly Arg Asp Glu Leu Phe Arg Thr Val Gly Trp Phe Thr Ser Ile His
 50 55 60
 Pro Val Val Leu Gly Ala Asp Arg Ser Val His Pro Glu Gln Tyr Leu
 65 70 75 80
 Ala Gln Ile Gly Ala Ala Leu Thr Ala Ala Pro Asp Gly Gly Val Gly
 85 90 95
 Phe Gly Ala Cys Arg Glu Phe Ser Pro Asp Ala Gly Leu Arg Thr Leu
 100 105 110
 Leu Arg Asp Leu Pro Pro Ala Leu Val Cys Phe Asn Tyr Tyr Gly Gln

125

Val ₁	Arg	Thr	Val	Arg ₅	Thr	Leu	Leu	Ile	Asp ₁₀	Asn	Tyr	Asp	Ser	Phe ₁₅	Thr
Tyr	Asn	Leu	Phe ₂₀	Gln	Met	Leu	Ala	Glu ₂₅	Val	Asn	Gly	Ala	Ala ₃₀	Pro	Leu
Val	Val	Arg ₃₅	Asn	Asp	Asp	Thr	Arg ₄₀	Thr	Trp	Gln	Ala	Leu ₄₅	Ala	Pro	Gly
Asp	Phe ₅₀	Asp	Asn	Val	Val	Val ₅₅	Ser	Pro	Gly	Pro	Gly ₆₀	His	Pro	Ala	Thr
Asp ₆₅	Thr	Asp	Leu	Gly	Leu ₇₀	Ser	Arg	Arg	Val	Ile ₇₅	Thr	Glu	Trp	Asp	Leu ₈₀
Pro	Leu	Leu	Gly	Val ₈₅	Cys	Leu	Gly	His	Gln ₉₀	Ala	Leu	Cys	Leu	Leu ₉₅	Ala
Gly	Ala	Ala	Val ₁₀₀	Val	His	Ala	Pro	Glu ₁₀₅	Pro	Phe	His	Gly	Arg ₁₁₀	Thr	Ser
Asp	Ile	Arg ₁₁₅	His	Asp	Gly	Gln	Gly ₁₂₀	Leu	Phe	Ala	Asn	Ile ₁₂₅	Pro	Ser	Pro
Leu	Thr ₁₃₀	Val	Val	Arg	Tyr	His ₁₃₅	Ser	Leu	Thr	Val	Arg ₁₄₀	Gln	Leu	Pro	Ala
Asp ₁₄₅	Leu	Arg	Ala	Thr	Ala ₁₅₀	His	Thr	Ala	Asp	Gly ₁₅₅	Gln	Leu	Met	Ala	Val ₁₆₀
Ala	His	Arg	His	Leu	Pro	Arg	Phe	Gly	Val	Gln	Phe	His	Pro	Glu	Ser

Page 36

165
 Ile Ser Ser Glu His Gly His Arg Met Leu Ala Asn Phe Arg Asp Leu
 180 185 190
 Ser Leu Arg
 195

(2) INFORMATION FOR SEQ ID NO: 25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 292 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 25:

Val Thr Ala Ala Ala Pro Thr Leu Ala
 1 5
 Gln Ala Leu Asp Glu Ala Thr Gly Gln Leu Thr Gly Ala Gly Ile Thr
 10 15 20 25
 Ala Asp Ala Ala Arg Ala Asp Thr Arg Leu Leu Ala Ala His Ala Cys
 30 35 40
 Gln Val Ala Pro Gly Asp Leu Asp Thr Cys Leu Ala Gly Pro Val Pro
 45 50 55
 Pro Arg Phe Trp His Tyr Val Arg Arg Arg Leu Thr Arg Glu Pro Ala
 60 65 70
 Glu Arg Ile Val Gly His Ala Tyr Phe Met Gly His Arg Phe Asp Leu
 75 80 85
 Ala Pro Gly Val Phe Val Pro Lys Pro Glu Thr Glu Glu Ile Thr Arg
 90 95 100 105
 Asp Ala Ile Ala Arg Leu Glu Ala Leu Val Arg Arg Gly Thr Thr Ala
 110 115 120
 Pro Leu Val Val Asp Leu Cys Ala Gly Pro Gly Thr Met Ala Val Thr
 125 130 135
 Leu Ala Arg His Val Pro Ala Ala Arg Val Leu Gly Ile Glu Leu Ser
 140 145 150
 Gln Ala Ala Ala Arg Ala Ala Arg Arg Asn Ala Arg Gly Thr Gly Ala
 155 160 165
 Arg Ile Val Gln Gly Asp Ala Arg Asp Ala Phe Pro Glu Leu Ser Gly
 170 175 180 185
 Thr Val Asp Leu Val Val Thr Asn Pro Pro Tyr Ile Pro Ile Gly Leu
 190 195 200
 Arg Thr Ser Ala Pro Glu Val Leu Glu His Asp Pro Pro Leu Ala Leu
 205 210 215
 Trp Ala Gly Glu Glu Gly Leu Gly Met Ile Arg Ala Met Glu Arg Thr

220
Ala Ala Arg Leu Leu Ala Pro Gly Gly Val Leu Leu Leu Glu His Gly
235 240 245
Ser Tyr Gln Leu Ala Ser Val Pro Ala Leu Phe Arg Ala Thr Gly Arg
250 255 260 265
Trp Ser His Ala Ser Ser Arg Pro Thr Cys Asn Asp Gly Cys Leu Thr
270 275 280
Ala Val Arg Asn His Thr Cys Ala Pro Pro Ala
285 290

(2) INFORMATION FOR SEQ ID NO: 26:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 75 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 26:

Ile Phe Glu His Lys Thr Val Ala Gln Leu Ala Pro Val Ala Glu
1 5 10 15
Thr Leu Ala Asp Thr Thr Arg Glu Glu Pro Ala Ala Val Ala Ala Thr
20 25 30
Gly Asp Val Pro Leu Thr Pro Ile Met His Trp Leu Arg Glu Arg Gly
35 40 45
Gly Pro Val Asp Ala Phe Ser Gln Thr Met Ala Val Thr Val Pro Ala
50 55 60
Gly Leu Asp Arg Glu Arg Leu Val Ala Ala Leu Gln
65 70 75

(2) INFORMATION FOR SEQ ID NO: 27:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 82 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 27:

Leu Glu Tyr Asp Thr Ala Leu Tyr Glu Arg Ala Thr Ala Glu Ala Leu
1 5 10 15
Thr Gly Arg Leu Leu Arg Leu Leu Asp Ala Val Val Thr Asp Pro Gln
20 25 30
Ala Pro Val Gly Ser His Asp Leu Leu Glu Glu Ala Glu His Ala Arg
35 40 45
Leu Ala Ala Phe Asn Asp Thr Ala Arg Pro Val Pro Arg Ala Gly Leu

50 55 60
Ala Glu Leu Phe Thr Ala Gln Ala Arg Arg Thr Ala Asp Ala Val Ala
65 70 75 80
Val Val

(2) INFORMATION FOR SEQ ID NO: 28:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 63 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 28:

Met Pro Pro Val Thr Pro Tyr Arg Ala Tyr Leu Ala His Leu Ala
1 5 10 15
Gly Arg Asp Asp Asp Ala Ala Arg Ala Ala Trp Arg Thr Ala Leu Ala
20 25 30
Asp Leu Glu Glu Pro Ser Leu Val Ala Gly Ala Gly Ala Gly Arg Gly
35 40 45
Ala Ala Asp Gly Ser Ala Leu Pro Gly Gln Ile Pro Gly Tyr Arg Ala
50 55 60

(2) INFORMATION FOR SEQ ID NO: 29:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 158 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 29:

Leu Gln Val Glu Gly Arg Pro Ala His Leu Glu Leu Pro Cys Asp His
1 5 10 15
Pro Arg Pro Ala Val Ala Thr His Arg Gly Ala Thr Val Pro Phe His
20 25 30
Ile Asp Ala Gly Leu His Glu Lys Leu Thr Ala Leu Ser Lys Ala Cys
35 40 45
Asp Ser Ser Leu Phe Met Val Leu Gln Ala Ala Val Ala Ala Leu Leu
50 55 60
Thr Arg His Gly Ala Gly Thr Asp Ile Pro Val Gly Ser Pro Val Ala
65 70 75 80
Gly Arg Thr Asp Asp Ala Leu Asp Asp Leu Val Gly Phe Phe Val Asn
85 90 95
Thr Leu Val Leu Arg Thr Asp Thr Ser Gly Asp Pro Thr Phe Arg Glu
100 105 110

03806.0054-04.seqlst

Leu Val Ala Arg Val Arg Gln Phe Asp Leu Ala Ala Tyr Thr His Gln
 115 120 125
 Asp Met Pro Phe Glu Lys Leu Val Glu Glu Val Asn Pro Glu Arg Ser
 130 135 140
 Leu Ala Arg Asn Pro Leu Phe Gln Val Val Leu Ala Leu Gln
 145 150 155

(2) INFORMATION FOR SEQ ID NO: 30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 161 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 30:

Met Pro Arg Ser Leu Asp Leu Tyr Val Ala Leu Leu Ala Val Leu
 1 5 10 15
 Lys Thr Gly Ala Ala Tyr Leu Pro Val Asp Ile Ser Tyr Pro Ala Glu
 20 25 30
 Arg Ile Ala Phe Met Ile Glu Asp Ala Arg Pro Val Thr Val Leu Asp
 35 40 45
 Arg Leu Pro Asp Asp Leu Gly Ala Tyr Arg Asp Thr Asp Leu Thr Asp
 50 55 60
 Ala Asp Arg Thr Ala Pro Leu Arg Pro Glu His Pro Ala Tyr Val Ile
 65 70 75
 His Thr Ser Gly Ser Thr Gly Thr Pro Lys Ala Val Val Met Pro His
 80 85 90 95
 Ala Gly Leu Val Asn Leu Leu Thr Trp His Ala Arg Arg Phe Pro Gly
 100 105 110
 Gly Thr Gly Val Arg Thr Ala Gln Phe Thr Ala Ile Gly Phe Asp Phe
 115 120 125
 Ser Val Gln Glu Ile Leu Ser Pro Leu Val Met Gly Lys Thr Leu Ala
 130 135 140
 Val Pro Ser Glu Glu Val Arg His Ser Ala Glu Leu Leu Ala Gly Trp
 145 150 155
 Leu Glu
 160

(2) INFORMATION FOR SEQ ID NO: 31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 97 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 31:

Leu Gln Ala Glu Gly Ala Glu Val Ser Leu Leu Ala Val Leu Asp Gly
 1 5 10 15
 Tyr Pro Asp Ala Tyr Asp Gly Thr Glu His Glu Val Gly Glu Glu Gln
 20 25 30
 Val Leu Ala Ile Leu Leu Asn Ala Ala Gly Val Asp Arg Ala Gln Ala
 35 40 45
 Phe Gly Asp Ala Pro Leu Gln Arg Ala Ala Val Leu Glu Lys Leu Arg
 50 55 60
 Asp Ser Gly Ser Ala Leu Gly Asn Leu Asp Asp Asp Ala Val Gly Arg
 65 70 75 80
 Met Val Thr Val Phe Leu Asn Asn Thr Arg Leu Ile Gln Asn Phe Arg
 85 90 95

Pro

(2) INFORMATION FOR SEQ ID NO: 32:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 23 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 32:

ATC GAY TTY CCN TAY CTS CCS GG 23

(2) INFORMATION FOR SEQ ID NO: 33:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 27 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 33:

TTC GAC GAY GAY GCN TTC GTS CAY GAC 27

(2) INFORMATION FOR SEQ ID NO: 34:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 base pairs
 - (B) TYPE: nucleic acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 34:

GTS CCS TGG CCS GAG GAC GTS GCS GCS AAG TAC 33

(2) INFORMATION FOR SEQ ID NO: 35:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 44 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 35:

GAG GTS GAG GGS CAC CTS GGS GCS CAC CCS GAC GTS CAG CAG GC 44

(2) INFORMATION FOR SEQ ID NO: 36:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 20 amino acids
 (B) TYPE: amino acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 36:

Val Pro Ala Ala Phe Val Pro Leu Asp Ala Leu Pro Leu Thr Gly Asn
 1 5 10 15
 Gly Val Leu Asp
 20

(2) INFORMATION FOR SEQ ID NO: 37:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 37:

GCS GCS TTC AAC GAC ACS GCS CGS CC 26

(2) INFORMATION FOR SEQ ID NO: 38:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 26 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 38:

TTC GTS CCS CTS GAC GCS CTS CCS CT 26

(2) INFORMATION FOR SEQ ID NO: 39:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 21 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 39:

GTS ACS CCS TAC CGS GCS TAC

21

(2) INFORMATION FOR SEQ ID NO: 40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 40:

ACS CGB CTS ATC CAG AAC TTC CGB CC

26

(2) INFORMATION FOR SEQ ID NO: 41:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 26 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 41:

TTC CGS GAC GCS ATG GCS CAG CTS GC

26

(2) INFORMATION FOR SEQ ID NO: 42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 38 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 42:

TTC GCS GGS GGS GAG TTC GCS GCS TGG GAC GGC ACC GG

38

(2) INFORMATION FOR SEQ ID NO: 43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 43:

GAC CCS GCS AAG CCS CCS CTS CTS TGG TAC CG

32